

Univerzita Karlova v Praze

Filozofická Fakulta

Bakalářská práce

2014

Helena Salačová

Univerzita Karlova v Praze
Filozofická fakulta
Ústav germánských studií

Bakalářská práce

Helena Salačová

Nahe Synonyme korpuslinguistisch betrachtet – Verben im Fokus

Near Synonyms from a Corpus Linguistic Perspective – Verbs in Focus

Korpusový výzkum blízkých synonym – slovesa

2014

Vedoucí práce:
Mgr. Martin Šemelík, Ph.D.

Prohlášení:

Prohlašuji, že jsem tuto bakalářskou práci vypracovala samostatně a výhradně s použitím citovaných pramenů, literatury a dalších odborných zdrojů.

V Praze, dne 11. Mai 2014

.....

Jméno a příjmení

Klíčová slova

současný německý jazyk – slovní zásoba – synonymie – korpusová lingvistika – kookurenční profily

Key words:

contemporary German – lexicon – synonyms – corpus linguistics – co-occurrence profiles

 Schlüsselwörter:

Gegenwartssprache – Wortschatz – Synonymie – Korpuslinguistik - Kookkurrenzprofile

Abstrakt

Tato bakalářská práce pojednává o blízké synonymii německých sloves, jejichž užití je pro nerodilé mluvčí často velmi problematické. Významy a shodné či rozdílné kontexty synonym (anfangen – beginnen – starten; herstellen – erzeugen – produzieren) jsou zkoumány pomocí Německého referenčního korpusu DeReKo, resp. kookurenční databanky CCDB. První část práce představuje problematiku synonymie; druhá, empirická část je věnována analýze CNS-map kontrastujících kookurenční profily vybraných sloves. Prostřednictvím těchto map je možné určit, v čem se daná synonyma liší či shodují.

Abstract

This bachelor thesis is concerned with the near synonymy of German verbs whose usage is oft very problematic for non-native speakers. Word meaning and identical or different context of synonyms (anfangen – beginnen – starten; herstellen – erzeugen – produzieren) are researched using the German reference corpus DeReKo, more specifically the cooccurrence databank CCDB. The first part of the thesis is a brief introduction to the vast field of synonymy; the other, empirical part is dedicated to analysis of CNS-maps contrasting co-occurrence profiles selected verbs. Using these maps we are enabled to determine the different similarities in the usage of these verbs.

Abstrakt

Diese Bachelorarbeit behandelt das Thema nahe Synonymie der deutschen Verben, deren Verwendung für Nichtmuttersprachler oft sehr problematisch ist. Die Bedeutungen und identischen bzw. unterschiedlichen Kontexte der Synonyme (anfangen – beginnen – starten; herstellen – erzeugen – produzieren) werden mithilfe des Deutschen Referenzkorpus DeReKo, bzw. der Kookkurrenzdatenbank CCDB untersucht. Im ersten Teil der Arbeit wird die Problematik der Synonymie dargestellt; der zweite, empirische Teil ist den Analysen von Merkmalskarten gewidmet, welche die Kookkurrenzprofile der ausgewählten Wörter kontrastieren. Anhand dieser Merkmalskarten ist festzustellen, wodurch sich die Synonyme unterscheiden bzw. worin sie übereinstimmen.

Danksagung

An dieser Stelle möchte ich mich bei Herrn Mgr. Martin Šemelik, Ph.D für seine Betreuung meiner Bachelor-Arbeit bedanken. Besonders schätze ich seine fachlichen und inspirativen Anregungen, Hilfsbereitschaft und ständige Motivierung.

Obsah

1. Vorwort	8
2. Synonymie.....	10
2.1. Definition der Synonyme bzw. Synonymie	10
2.2. Klassifikation	12
2.3. Aufgabe des Kontextes bei der Synonymie.....	13
2.4. Funktion der Synonymie	15
2.5. Nahe Synonymie	16
3. Korpuslinguistik.....	17
3.1. Zur CCDB	19
3.2. Erklärung der terminologischen Begriffe	20
3.2.1. Kookkurrenz	20
3.2.2. Kookkurrenzanalyse	21
3.2.3. Kookkurrenzprofil (KP)	21
3.2.4. Merkmalskarte	23
4. Empirischer Teil – Analyse der synonymischen Reihen.....	29
4.1. Analyse der Synonymreihe: <i>ANFANGEN – BEGINNEN – STARTEN</i>	30
4.1.1. Analyse <i>anfangen – beginnen</i>	31
4.1.2. Analyse <i>beginnen-starten</i>	37
4.1.3. Analyse <i>anfangen – starten</i>	40
4.2. Analyse der Synonymreihe: <i>HERSTELLEN – ERZEUGEN – PRODUZIEREN</i>	44
4.2.1. Analyse <i>herstellen – erzeugen</i>	45
4.2.2. Analyse <i>herstellen – produzieren</i>	49
4.2.3. Analyse <i>erzeugen – produzieren</i>	51
5. Praktische Anwendung der CCDB.....	55
5.1. Lexikografie	55
5.2. Anwendung im DaF-Unterricht	57
6. Arbeitsblatt zu MK 1 (CNS <i>anfangen – beginnen</i>): Ein Entwurf.....	60
7. Fazit.....	64
8. Literaturverzeichnis.....	66
9. Im Text verwendete Abkürzungen	68

1. Vorwort

Die vorliegende Bachelorarbeit behandelt ein lexikalisches Thema, und zwar die nahe Synonymie in der deutschen Sprache. Die Synonymie bereichert den Wortschatz einer Sprache und die Muttersprachler der jeweiligen Sprache gebrauchen sie automatisch zur Schattierung der Aussage. Nichtmuttersprachlern hingegen bereitet die praktische Anwendung oft Schwierigkeiten, denn sie können nicht immer die häufig sehr feinen Unterschiede zwischen inhaltlich nahen Synonymen erkennen. Da ich selbst Deutsch als Fremdsprache lerne und studiere und auch später unterrichten will, möchte ich mich mit solchen deutschen Synonymen beschäftigen, die semantisch sehr nah beieinander liegen und die lediglich über ein einziges Äquivalent in der tschechischen Sprache verfügen. Wenn die Übersetzung der Synonyme nur mit einem Wort ins Tschechische möglich ist, können die tschechischen Sprecher keine offensichtlichen Differenzen zwischen den Lexemen erkennen. Das ist nicht nur ein Problem der tschechischen Sprecher, sondern auch der zweisprachigen Wörterbücher, die sehr oft ungenügende Informationen zur Verfügung stellen. Dies hat mich dazu angespornt, solche deutschen synonymischen Lexeme zur Analyse auszuwählen, deren Verwendung ich persönlich für problematisch halte und bei denen ich selbst Unklarheiten empfinde. Das Ziel dieser Arbeit ist die Analyse synonymischer Wortpaare, die ich aufbauend auf den der Sprachdatenbank zugrundeliegenden lexikalischen Merkmalskarten untersuche. Die ausgewählten deutschen Synonyme werden nicht nur miteinander, sondern auch mit einem entsprechenden Lehnwort verglichen. Im Fokus der Forschung stehen zwei synonymische Reihen: *anfangen* – *beginnen* und *starten*; *herstellen* – *erzeugen* und *produzieren*. Mein Vorhaben war es, ihre Verwendungsweisen und Differenzen bzw. Übereinstimmungen festzustellen.

Zunächst stelle ich die Problematik der Synonymie vor (Kap. 1). In diesem Kapitel gehe ich auf Definition der Synonymie, ihre Funktion und problematische Klassifizierung ein. Weiter wird diskutiert, inwieweit der Kontext einen Einfluss auf Synonymie hat. Ein Teil des Kapitels wird der nahen Synonymie gewidmet, denn sie ist das Hauptthema meiner Arbeit.

Im zweiten Kapitel behandle ich die Problematik der Korpuslinguistik als Methodologie, die bei der linguistischen Forschung eine wichtige Rolle spielt. Als

Nächstes wird die Kookkurrenzdatenbank CCDB vorgestellt, die einen Bestandteil des Deutschen Referenzkorpus bildet und mit deren Hilfe die Analysen durchgeführt werden. Um die Funktion der CCDB zu verdeutlichen, erläutere ich die Hauptbegriffe wie Kookkurrenz, Kookkurrenzanalyse und Kookkurrenzprofil. Die CCDB verfügt über spezielle Tools, mit deren Hilfe die Synonyme analysiert werden können. Es handelt sich um die sog. Merkmalskarten, die eine Visualisierung der Lexeme ermöglichen (SOM-Merkmalskarten) und auch zwei Lexeme und deren Kookkurrenzprofile miteinander kontrastieren können (CNS-Merkmalskarten).

Im empirischen Teil der Arbeit werden die von mir ausgewählten Synonyme untersucht, und zwar mit dem Einsatz der korpuslinguistischen Methode (mit der CCDB). Mithilfe der Merkmalskarten versuche ich, die grundlegenden Unterschiede und Gemeinsamkeiten der Synonyme zu ermitteln, besonders ihre kontextuellen Verwendungsweisen. Die Ergebnisse vergleiche ich gleichzeitig mit Angaben aus dem Duden-Universalwörterbuch.

Das letzte Kapitel bietet einen Einblick in den Bereich der praktischen Anwendung der Merkmalskarten. Es wird darauf aufmerksam gemacht, dass Wörterbücher nicht immer vollständige Informationen liefern. Deshalb ist die CCDB ein nützliches Hilfsmittel, das genauere Auskünfte über Lexemgebrauch und -bedeutung geben kann. Zuletzt widme ich mich der Didaktik und einem potenziellen Gebrauch von lexikalischen Merkmalskarten im DaF-Unterricht. Als Beispiel präsentiere ich einen Entwurf meines eigenen Arbeitsblatts für den DaF-Unterricht.

2. Synonymie

Die synonymische Beziehung zwischen Lexemen ist zwar allgemein bekannt (meistens als semantische Beziehung zwischen Lexemen ähnlicher oder gleicher Bedeutung verstanden), aber in Bezug auf mehrere Aspekte problematisch. Die unterschiedlichen Auffassungen verschiedener Linguisten beruhen auf der Definition und der damit verbundenen Klassifikation der Synonyme. Dieses Kapitel dient als Einführung in diese sehr weite Problematik und beinhaltet Ansichten von einigen Lexikologen, die sich mit diesem Thema beschäftigen bzw. beschäftigt haben. Außerdem werden Funktionen der Synonyme angesprochen und es wird auch kurz beschrieben, wie sie sich erforschen lassen. Schließlich wird der Begriff „nahe Synonymie“ ausführlicher diskutiert, weil ich mich mit ausgewählten Synonymreihen und ihrer Erforschung in den nächsten Kapiteln beschäftige.

2.1. Definition der Synonyme bzw. Synonymie

„Synonyme sind sprachliche Einheiten oder Strukturen, die sich formal unterscheiden, aber ähnliche oder gleiche Bedeutung haben und deshalb im Kern der Bedeutung übereinstimmen. Synonymie ist die Bezeichnung für die Beziehung zwischen Synonymen.“ (Schippan 2002: S. 206)

Nach dieser allgemeinen Definition ist Synonymie eine lexisch-semantische Beziehung zwischen Lexemen (Synonymen), die miteinander sinnverwandt sind. Synonyme sind Lexeme des gleichen Denotats (Objektes, auf das sich das Lexem bezieht) und sie verfügen über einen Kern gleicher Bedeutungselemente. Sie können sich durch denotative Seme oder konnotative Merkmale unterscheiden (ibidem: S. 207). Eine solche Auffassung ist der strukturellen Semantik zuzuordnen (vgl. Marková 2012a: S. 15ff).

Die Definition von Filipec ist noch breiter:

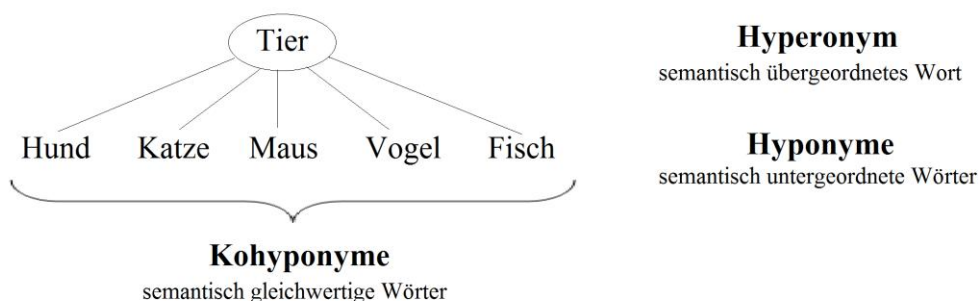
„Synonyma jsou lexikální jednotky [...] se stejnou nebo částečně odlišnou oblastí kontextového užití [...]“ (Filipec 1961: S. 203)

[Synonyme sind lexikalische Einheiten mit dem gleichen oder teilweise verschiedenen Bereich des kontextuellen Gebrauchs. (Meine Übersetzung, H.S.)]

Aus seiner Definition geht hervor, dass die lexikalischen Einheiten ihre synonymischen Beziehungen nicht nur an sich selbst zeigen, sondern dass Synonyme auch mit Bezugnahme auf Ähnlichkeit ihrer Kontexte entstehen, in denen sie vorkommen. Daraus könnte man logisch schlussfolgern, dass sich Synonyme in ihren Kontexten austauschen lassen, das gilt jedoch nicht für alle Fälle. Einige Synonyme sind in denselben Kontexten substituierbar, andere hingegen haben eigene spezifische Kontexte (vgl. Filipec 1961: S. 188).

Erklärungen, warum zwei bedeutungsidentische Lexeme in der Sprache so selten vorkommen, geben Linke et al. (2004: S. 172) an. Der erste Grund ist, dass Sprachen zur Simplizität tendieren, zum sog. Prinzip der Ökonomie. Zwei identische Wörter wären deshalb überflüssig. Der zweite Grund wird bezeichnet als das Prinzip der größtmöglichen Differenziertheit: *„die Ausdrücke werden mit unterschiedlichen Bedeutungen oder Bedeutungsnuancen aufgeladen und damit verschieden verwendet“* (ibidem).

Synonyme gehören aufgrund gemeinsamer semantischer Merkmale zu paradigmatischen Beziehungen. Das bedeutet, dass die lexikalischen Einheiten in gleichen Kontexten und syntaktischen Strukturen auftreten können. Damit diese Aussage gelten kann, müssen sie auch gleicher Wortart sein. Neben Synonymen werden noch Hyperonyme, Hyponyme, Kohyponyme und Antonyme den paradigmatischen Beziehungen zugeordnet (vgl. Schippan 2002: S. 202ff).



(meine Grafik, H. S.)

2.2. Klassifikation

Synonyme lassen sich klassifizieren nach dem Grad der semantischen Nähe, der jedoch sehr schwer zu ermitteln ist. Es gibt mehrere Typen von Klassifikationen, denn die Linguisten sind sich über eine exakte Definition der Synonymie und ihrer Kategorien nicht einig. Die Grenzen zwischen zwei Synonymen sind nicht präzise festzustellen. Manche Lexikologen vertreten die Auffassung, dass Synonymie lediglich partiell vorkommt, d. h., dass synonymische Lexeme nicht völlig frei austauschbar sein können ohne die Bedeutung zu beeinflussen (vgl. Schippan 2002: S. 213). Dieser Ansicht stimme ich zu.

Z. B. in Filipec/Čermák (1985: S. 133ff) werden viele Typen der Synonymie angeführt und zwar nach Spezifikation, Intensität, Pragmatik, Konnotation u. a. Von den vielfältigen Typen sind dann zwei Hauptgruppen zu unterscheiden: identische und partielle Synonymie.

Genauso wird die Klassifikation der Synonymie traditionell betrachtet. Filipec (1961), Linke et al. (2004) u. a. unterteilen Synonymie in **absolute** (bzw. „identische“ nach Filipec/Čermák 1985 oder „reine“ nach Ullmann 1967) und **partielle**. Cruse (2011) hingegen gibt drei Gruppen von synonymischen Beziehungen an: absolute, propositionale und nahe Synonymie, wobei er die absolute Synonymie eher theoretisch meint (praktisch kommt sie seiner Meinung nach nicht vor. Diese drei Typen sind wie folgt zu verstehen:

Absolute Synonymie nach Cruse (2011: S. 142):

„Absolute synonyms [are] words which are mutually substitutable in all contexts without change of normality.“

„ [...] absolute synonyms are vanishingly rare, and do not form a significant feature of natural vocabularies.“ (ibidem: S. 143)

Mit dieser Definition weicht Cruse nicht von anderen Linguisten ab. Er erwähnt aber eine Besonderheit in Form der **propositionellen Synonymie** (ibidem: S. 143):

„Two sentences which differ only in that one has one member of a pair of propositional synonyms where the other has the other member of the pair are mutually entailing.“

Er erklärt diesen Typ der Synonymie am Beispiel der englischen Wörter *fiddle* und *violin*. Beide Wörter bedeuten „Geige“, aber ihr Gebrauch hängt von der Einstellung des Sprechers ab (*fiddle* für einen Berufsmusiker bzw. Violinisten, *violin* für einen Laien). Die Substitution von diesen Synonymen wird also nur in bestimmten Situationskontexten realisiert.

Marková (2012a: S. 26) fügt zur Definition von Cruse hinzu, dass man propositionelle Synonyme nur innerhalb von Sätzen bestimmen kann. Das heißt, sie sind Teile eines Syntagmas, in dem sie entstehen und in dem sie durcheinander substituiert werden können.

Der letzte Grad der Klassifikation nach Cruse ist die sog. **nahe Synonymie**, die nicht so einfach von der propositionellen zu unterscheiden ist. Eine typische Eigenschaft der nahen Synonymie ist das Maß des Kontrasts. Die Kontrastierung ist zwar nicht die primäre Funktion der Synonymie allgemein, bei der nahen Synonymie spielt sie aber eine Rolle. Dieser Kontrast wird oft nur implizit ausgedrückt, in bestimmten Umgebungen kommt er jedoch explizit vor, z. B.:

„*He was killed, but I can assure you he was NOT murdered, madam.*“ (Beispiel von Cruse 2011: S. 145)

2.3. Aufgabe des Kontextes bei der Synonymie

Wie schon erwähnt, darf der Kontext bei der Untersuchung von lexikalischen Beziehungen nicht unterlassen werden. Um nämlich festzustellen, ob zwei Wörter Synonyme sind, müssen sie in einer bestimmten Situation und einem bestimmten Kontext die gleiche Bedeutung haben. Aus diesem Grunde sind Lexeme nur in eigenen synonymischen Kontexten wirklich Synonyme (vgl. Filipec/Čermák 1985: S. 133).

Beispiel:

Mit ihm ist heute nichts anzufangen. (Er ist heute nicht in Form, nicht ansprechbar; s. Tabelle 1; DUW) In diesem Kontext hat das Verb *anfangen* eine spezifische Bedeutung und kann durch *beginnen* nicht substituiert werden. Im folgenden Satz hingegen sind beide Verben zu verwenden, d. h. sie sind synonymisch:

Ich habe mit der Arbeit angefangen/begonnen. (vgl. DUW, Tabelle 1)

Nach Schmidt (2012: S. 125) wird anhand von bisherigen Forschungsarbeiten die Hypothese unterstützt, „dass zwischen dem Verhältnis zweier synonymischer Lexeme und der Ähnlichkeit der Kontexte, in denen diese Lexeme erscheinen, ein Zusammenhang besteht. Demnach ist es [...] nötig, die Aufmerksamkeit nicht nur auf das Lexem an sich zu richten, sondern auch auf die Umgebung, in der es vorkommt, bzw. auf lexikalische Einheiten, die in dieser Umgebung signifikant häufig erscheinen.“

Damit wird die schon fünfzig Jahre alte Theorie von Filipec (1961: S. 188) untermauert:

„Studium oblasti kontextového užití se musí zabývat i synonymika, neboť jednotlivá synonyma se vyskytují jednak v stejných, jednak v různých kontextech a z hlediska lexikologie jde pak o vztah celých oblastí, které musíme u jednotlivých lexikálních synonym porovnávat.“

[Mit dem Studium des Bereichs der kontextuellen Verwendung muss sich auch die Synonymie beschäftigen, denn einzelne Synonyme kommen sowohl in gleichen als auch in unterschiedlichen Kontexten vor und unter dem Aspekt der Lexikologie geht es dann um die Beziehung von ganzen Bereichen, die bei einzelnen lexikalischen Synonymen zu vergleichen sind. (meine Übersetzung, H. S.)]

Die Wichtigkeit des Kontextes betont ebenso Storjohan (2006: S. 2), indem sie darauf aufmerksam macht, dass zwischen Lexemen keine paradigmatischen Sinnrelationen existieren, sondern dass diese erst in kontextuellen Umgebungen entstehen. Da zwischen Synonymen eine paradigmatische Beziehung besteht, gilt die Aussage von Storjohan auch für sie.

Bedeutungen und Kontextanwendungen zweier Wörter können einander ähneln. Wichtig ist dabei, dass der Grad der Bedeutungs- und Kontextähnlichkeit nicht identisch sein muss. Denn auch Antonyme können über ähnliche Kontexte verfügen (vgl. Filipec: S. 189).

Die Ermittlung von Synonymen kann mithilfe von Substitutions- oder Lückentests durchgeführt werden. Damit erweist sich die Relevanz der Verbindung mit dem Kontext, obwohl es sich in diesem Zusammenhang lediglich um einen Minimalkontext handelt.

Bei Substitutionstests werden Lexeme (Synonyme) in den Minimalkontext eingesetzt und dann werden Übereinstimmungen oder Differenzen festgestellt. Die Substitutionstests ermöglichen Merkmale und Konnotationen zu verdeutlichen und sie verhelfen denotative und konnotative Seme zu ermitteln und den Grad der Synonymie zu bestimmen. Dieser Typ der Untersuchung gilt jedoch als veraltet und wird z. B. nur als eine Hilfsmethode im Deutsch-Unterricht angewendet. Die Ersatzprobe wird so realisiert, dass die syntaktische Struktur nicht verletzt wird und die Wahrheit der Aussage nicht verändert werden darf (vgl. Schippan 2002: S. 208 und Storjohan 2006: S. 5).

Es sind drei Faktoren zu unterscheiden, nach denen Synonymie untersucht werden kann und die mit dem Kontext zusammenhängen (Schmidt 2012: S. 124):

1. Übereinstimmung in der Bedeutung – *Intension*
2. Übereinstimmung in der *Distribution* – Umgebung, Kontexte, Texttypen
3. Übereinstimmung im Bedeutungsumfang – *Extension* (d. h. Kombinierbarkeit eines Lexems mit anderen lexikalischen Elementen, sog. Kollokabilität)

Bei einer solchen Untersuchung von lexikalischen Umgebungen und Kontexten verfügen Korpora über eine bedeutende Funktion. Insbesondere die Kookkurrenzdatenbank CCDB dient dazu, feine Bedeutungsnuancen der synonymischen Lexeme anhand vom topografischen Modell der semantischen Proximität (Nähe) zu enthüllen (s. u.).

2.4. Funktion der Synonymie

Im üblichen Sprachgebrauch ist die Wahl der Synonymie nicht willkürlich. Man wird sich ihrer Existenz meistens erst bei der Formulierung einer wichtigen Äußerung bewusst. Synonyme helfen nämlich einen Gedanken zu konkretisieren und deshalb müssen sie sorgfältig ausgewählt werden (vgl. Schippan 2002: S. 209).

Synonyme, die im gleichen Text auftreten, besitzen verschiedene Merkmale. Ihre Merkmale erweisen sich erst im Kontext. Mit Synonymen können Aspektbetonung, Steigerung, Erläuterung oder Emotionalität und Wertung ausgedrückt werden (ibidem).

2.5. Nahe Synonymie

Nahe Synonymie ist ein Typ der Synonymie, der als Plesionymie bezeichnet wird (vgl. Storjohan 2006: S. 11). Diese Beziehung zwischen Lexemen stellt eine dichte semantische Nähe dar, d. h. die Bedeutungen beider Lexeme ähneln einander. Diese Lexeme besitzen eine Menge gemeinsamer semantischer Merkmale. Die Abgrenzung, wann es sich um nahe Synonyme genau handelt und wie viel semantische Distanz solche Synonyme brauchen, ist schwer festzulegen. Plesionymie lässt sich so verstehen, dass die Kontraste zwischen Lexemen auch ohne Kontext evident sein müssen, d. h., ihre Unterschiede dürfen nicht kontextuell fokussiert werden. Deshalb ist zu konstatieren, dass Synonymwörterbücher über einen großen Teil solcher Angaben verfügen, die aus Plesionymen bestehen (vgl. ibidem). Storjohan führt konkrete Beispiele der nahen Synonyme an: *lachen* vs. *lächeln*, *laufen* vs. *rennen* oder *heiß* vs. *kochend* u. a. Dazu sind auch die weiteren verglichenen synonymischen Paare, die ich mithilfe eines Korpus analysiert habe, zu zählen (z. B. *anfangen* vs. *beginnen*, s. u.).

Da aber trotzdem Unterschiede zwischen nahen Synonymen existieren, muss der Sprecher wissen (bei Muttersprachlern meistens unwillkürlich), worin die Unterschiede zwischen ihnen bestehen. Erst dann kann er sie richtig verwenden. In Bezug auf Deutsch haben die Sprecher, deren Muttersprache eine andere Sprache als Deutsch ist, bei der Verwendung der nahen Synonyme oft Schwierigkeiten. Sie müssen die Unterschiede bewusst erlernen oder gezielt untersuchen, wie es in meiner Arbeit der Fall ist (s. die Analysen unten).

3. Korpuslinguistik

Bevor auf die Kookkurrenzdatenbank CCDB eingegangen wird, ist kurz die Auffassung der Korpuslinguistik zu thematisieren.

Die Korpuslinguistik ist eine Methode der Sprachanalyse:

„[...] it is an area which focuses upon a set of procedures, or methods, for studying language [...]” (McEnery, Hardie 2012: S. 1)

Zwar wird die Korpuslinguistik meistens als methodologischer Ansatz betrachtet, jedoch sind die Ansichten der Linguisten über diesen Zweig der Sprachwissenschaft unterschiedlich. Nicht alle stimmen der Auffassung „Korpuslinguistik als Methodologie“ zu, sondern sie betrachten sie als eine selbstständige wissenschaftliche Disziplin getrennt von anderen Bereichen der Linguistik (vgl. Perkuhn et al. 2012: S. 18; zur ausführlicher Erklärung siehe Tognini-Bonelli 2001). Korpuslinguistik als Disziplin beschäftigt sich mit dem Aufbau der Korpora, mit der Sprechdatenverarbeitung, und deren Speicherung und Auswertung.

In dieser Arbeit wird die Ansicht „Korpuslinguistik als Methodologie“ bevorzugt, denn der Aufbau von Korpora ist für mein Thema nicht relevant. Der korpuslinguistische Ansatz stützt sich auf empirische Beobachtungen von Sprechdaten und darauf beruht der Inhalt dieser Arbeit – der Vergleich von nahen Synonymen mithilfe eines Korpus.

Die Korpuslinguistik als Methodologie der Spracherforschung ist eine empirisch aufgebaute Wissenschaft, die anhand von Daten die Sprache strukturalistisch und deskriptiv beschreibt. Sie ermöglicht ein neues Studium der Sprache und ihrer Erscheinungen, vor allem auf der syntagmatischen Ebene. Die Richtung der Korpuslinguistik ist in der letzten Zeit besonders an Untersuchungen von Kollokationen orientiert. Korpuslinguisten arbeiten mit einer breiten Menge von Sprachdaten, weshalb diese Art der Wissenschaft im Rahmen der Linguistik sehr bedeutsam ist (vgl. Čermák 2011: S. 21ff).

Korpuslinguistische Forschung und Äußerungen müssen sich auf große Belegzahlen stützen. D. h. es dürfen nicht lediglich anhand weniger Belege grundlegende Theorien formuliert werden. Erkenntnisse sind erst dann zu gewinnen, wenn die Belege im Rahmen des Korpus als Gesamtheit verstanden werden, also wenn verschiedene

Sprachdaten ins Blickfeld genommen werden (vgl. Perkuhn, Belica 2006: S. 2). Wichtig ist zu erwähnen, dass die Korpusergebnisse nur positiv betrachtet werden können, denn nur wenn ein Phänomen nachgewiesen ist, existiert es auch. Umgekehrt gilt das jedoch nicht: Wenn eine Form im Korpus nicht nachzuweisen ist, muss es nicht unbedingt heißen (aber kann), dass sie z. B. in der gesprochenen Sprache nicht verwendet werden kann und so nicht akzeptabel ist. Aus diesem Grund ist keine Aussage aus der Absenz einer Form zu folgern. Weiter ist zu beachten, dass Korpora nicht selten grammatische Fehler beinhalten, was aber nur als Fakt über den Sprachgebrauch zu verstehen ist. Grammatische Fehler oder Formen, die der Standardsprache nicht entsprechen, weisen auf die Authentizität des Korpusinhaltes hin (vgl. ibidem: S. 4-5) und sie sind ein Teil des Sprachgebrauchs.

Korpora verfügen zwar über wertvolle authentische Sprachdaten, aber sie sind nicht in der Lage, die gesamte Sprache vollkommen zu umfassen. Der reale Sprachgebrauch wird deshalb mit Einschränkungen im Korpus präsentiert. Demzufolge besteht eine Asymmetrie zwischen den Korpora und dem Sprachgebrauch (vgl. Marková 2012a: S. 18). Dieser Aussage entspricht u. a. die Äußerung von Čermák:

„[...] nelze zjistit proporce jednotlivých typů promluv, natožpak je jasné definovat [...].“ (Čermák 2011: S. 16)

[Es ist nicht möglich, Proportionen einzelner Typen von Aussagen festzustellen, geschweige dann sie klar zu definieren. (meine Übersetzung, H. S.)]

Das Problem besteht darin, dass es nicht realisierbar ist, die gesprochene Sprache vollständig in die Korpora aufzunehmen. Erstens ist die Einordnung technisch nicht einfach, zweitens sind authentische Aufnahmen nicht problemlos zu gewinnen. Natürlich sind sie z. B. aus dem Rundfunk leicht zugänglich. Diese Daten haben jedoch in hohem Maß Merkmale eines geschriebenen Textes, und sind deshalb für die Untersuchung der gesprochenen Sprache nicht besonders geeignet. Bei der gesprochenen Alltagssprache ist die Aufnahme wesentlich schwieriger als bei geschriebenen Texten.

Aus diesem Grunde enthalten die Korpora vor allem Daten der geschriebenen Sprache. Beim Deutschen Referenzkorpus (DeReKo), das den Untersuchungen im Rahmen meiner Arbeit zugrunde liegt, handelt es sich ausschließlich um die

deutschgeschriebene Gegenwartssprache. Dieses Korpus verfügt über Zeitungsartikel, belletristische, wissenschaftliche und populärwissenschaftliche Texte (unter <http://www1.ids-mannheim.de/kl/projekte/korpora/>, Stand: 1/2014).

In der Korpuslinguistik wird die Sprache auf der Ebene der *Parole* untersucht, d. h. die korpuslinguistische Untersuchung ist auf Sprachgebrauch, realisierte Äußerungen und Texte bezogen. Ein Bestandteil der *Parole* ist jedoch auch das Sprachsystem, die sog. *Langue*, die in der *Parole* verborgen ist. Aus dieser Tatsache geht hervor, dass die *Langue*-Ebene im Korpus gleichfalls beinhaltet ist, denn die *Parole* wird mithilfe der *Langue* realisiert (vgl. Marková 2012a: S. 18).

Im Rahmen der Korpuslinguistik werden zwei Typen von Methodologien differenziert: *corpus-based* und *corpus-driven*. Als *corpus-based* wird ein Vorgang der Forschung bezeichnet, bei dem man anhand von Theorien und Thesen konkrete Beispiele oder Erkenntnisse nachträglich im Korpus auffindet (vgl. Perkuhn et al. 2012: S. 20). Dagegen funktioniert es bei der *corpus-driven*-Methodologie umgekehrt:

„Daten dienen nicht dazu, erst im Nachhinein Thesen oder Theorien zu bestätigen oder zu widerlegen. Sie stellen vielmehr den Ausgangspunkt dar, von dem aus Thesen abgeleitet und Theorien aufgestellt werden.“ (ibidem: S. 20)

In der Praxis werden diese zwei Methodologien in der Regel kombiniert, so dass der Forscher zu objektiveren Erkenntnissen kommen kann.

3.1. Zur CCDB

Die Kookkurrenzdatenbank CCDB ist eine Denk- und Experimentierplattform, die im Jahre 2001 am Institut für Deutsche Sprache (IDS) in Mannheim entstand. Die CCDB ermöglicht eine Weiterentwicklung von Methoden der Kookkurrenzanalyse von Lexemen und deren Forschung. Sie entwirft eine kohärente Methodik, welche die Sprachstrukturen systematisch enthüllt und welche es ermöglicht, die Ergebnisse zu interpretieren und theoretisch zu erklären. Sie geht von empirischen Daten aus, die im DeReKo enthalten sind. Die CCDB basiert auf der Menge von Wörtern, über die das DeReKo verfügt. Es sind mehr als 2,2 Milliarden Textwörter (dazu werden alle Wörter zugerechnet, die in den Texten im DeReKo eingegliedert sind; d. h. auch die, die sich ständig wiederholen und

deren Frequenz hoch ist) und 220 000 Stichwörter, d. h. dem Benutzer steht diese Menge verschiedener Wörter zur Verfügung. Die CCDB dient zur Suche nach Kookkurrenzen und wobei die zugrunde liegende Software automatisch ein Kookkurrenzprofil erstellt. Der Algorithmus, der die Kookkurrenzen im DeReKo sucht und auswertet, wird schon seit 1995 benutzt und ist ein Bestandteil des Recherchesystems COSMAS (vgl. Keibel/Belica 2007 und <http://corpora.ids-mannheim.de/ccdb/>, Stand 1/2014).

3.2. Erklärung der terminologischen Begriffe

Im folgenden Abschnitt werden Begriffe wie **Kookkurrenz**, **Kookkurrenzanalyse**, **Kookkurrenzprofil** und **Merkmalskarte** erklärt.

3.2.1. Kookkurrenz

Zu jedem Wort, das zu einem Text gehört, bilden seine benachbarten Wörter einen sprachlichen Kontext. Oft wird der Kontext so verstanden, dass er von jeweils fünf Wörtern auf der linken und rechten Seite vom Wort entsteht. So lässt sich der Kontext natürlich nicht generell betrachten. Die Größe des Kontextes hängt oft von den zu untersuchenden Erscheinungen ab (vgl. Perkuhn et al. 2012: S. 111).

Nach Bußmann versteht man unter der Kookkurrenz eine Umgebung, in der sich ein Wort zusammen mit anderen Wörtern befindet:

„[Kookkurrenz ist eine] Prägung, die das Miteinandervorkommen von sprachlichen Elementen verschiedener Klassen in Sätzen bezeichnet. Kookkurrenz eines Elements ist die Summe aller syntaktischen Umgebungen, in denen es vorkommen kann.“ (Bußmann: 2002)

Der Begriff „Kookkurrenz“ bedeutet also das rekurrente gemeinsame Vorkommen sprachlicher Einheiten (vgl. Perkuhn et al. 2012: S. 113). D. h. Kookkurrenz stellt ein wiederholtes Auftreten von Wörtern dar, die zwar semantisch nicht zusammen gehören oder miteinander verwandt sein müssen, die aber oft in ähnlichen Kontexten zusammen auftreten.

3.2.2. Kookkurrenzanalyse

Die Wortkombinationen lassen sich mittels eines mathematisch-statistischen Verfahrens analysieren – es wird die sog. Kookkurrenzanalyse abgebildet. Ihr Gegenstand ist es, die folgende Frage zu beantworten: In Verbindung mit welchen Wörtern kommt ein bestimmtes Wort am häufigsten vor? Eine Kookkurrenzanalyse wird deshalb auf der syntagmatischen Ebene durchgeführt.

Anders gesagt handelt es sich bei der Kookkurrenz um den sog. lokalen Kontext, in dem die einzelnen lexikalischen Einheiten gebraucht werden (Belica: 2007). Im Unterschied zum globalen Kontext umfasst der lokale Kontext (die Kookkurrenz) keine situativen Kontexte, die einen thematischen Bereich darstellen. (Jedoch ist es anhand der Analyse vom lokalen Kontext möglich, den globalen Kontext zu bestimmen.)

3.2.3. Kookkurrenzprofil (KP)

Was unter dem Begriff Kookkurrenzprofil zu verstehen ist, wird in der Definition von Belica erklärt:

„Die Gesamtheit aller quantitativen Ergebnisse der Kookkurrenzanalyse zu einem gegebenen Analyseobjekt (einem Lexem, einer Wortverbindung usw.) wird als Kookkurrenzprofil des Objektes bezeichnet und stellt – informell gesagt – ein Kondensat seines Gebrauchs dar. Es erfasst sowohl dominante Wortverbindungsstrukturen wie auch subtile Varianzphänomene im lokalen lexikalischen Kontext des analysierten Objektes, und bietet dadurch eine detaillierte Auskunft über die syntagmatische und paradigmatische Einbettung des Objekts im Sprachgebrauch aus präferenzrelationaler Sicht.“ (Belica: 2007)

Anders gesagt: Ein KP ist ein Komplex von Kookkurrenzen des analysierten Objektes. Es können entweder einzelne Lexeme oder ganze Wortverbindungen analysiert werden. Mithilfe des KP lässt sich das Lexem im Rahmen des Sprachgebrauchs näher bestimmen. Dabei wird einerseits die syntagmatische Einbettung berücksichtigt – d. h. wie sich das Lexem im lexikalischen Kontext verhält und welche Lexeme es umgeben, und andererseits die paradigmatische Einbettung – d. h. welche anderen Lexeme das Lexem in dieser Position vertreten können (es handelt sich z. B. um Synonyme oder Antonyme).

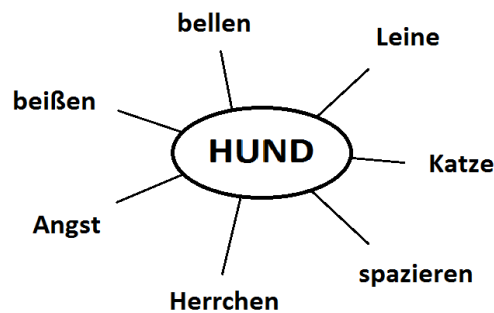


Abb. 1: Das Kookkurrenzprofil von *Hund*. Meine Grafik, H. S. (Die Wörter stammen aus der Kookkurrenzanalyse der CCDB.)

Dank der Analyse von Kookkurrenzprofilen lassen sich viele Aspekte und Nuancen der Verwendung von Lexemen ermitteln.

Um es noch klarer zu machen, führe ich ein vereinfachtes visualisiertes Beispiel an (Abb.1). Hier ist verdeutlicht, mit welchen Wörtern (Kookkurrenzpartnern) das Analyseobjekt *Hund* am häufigsten vorkommt. Die Kookkurrenzpartner bilden sein Kookkurrenzprofil.

Als weiteres Beispiel ist ein Ausschnitt aus der CCDB unter Abb. 2 zu sehen. Es handelt sich um einen Ausschnitt des Kookkurrenzprofils vom Verb *starten*, das zusammen mit *anfangen* und *beginnen* in Kapitel 4.1.2. und 4.1.3. analysiert wird.

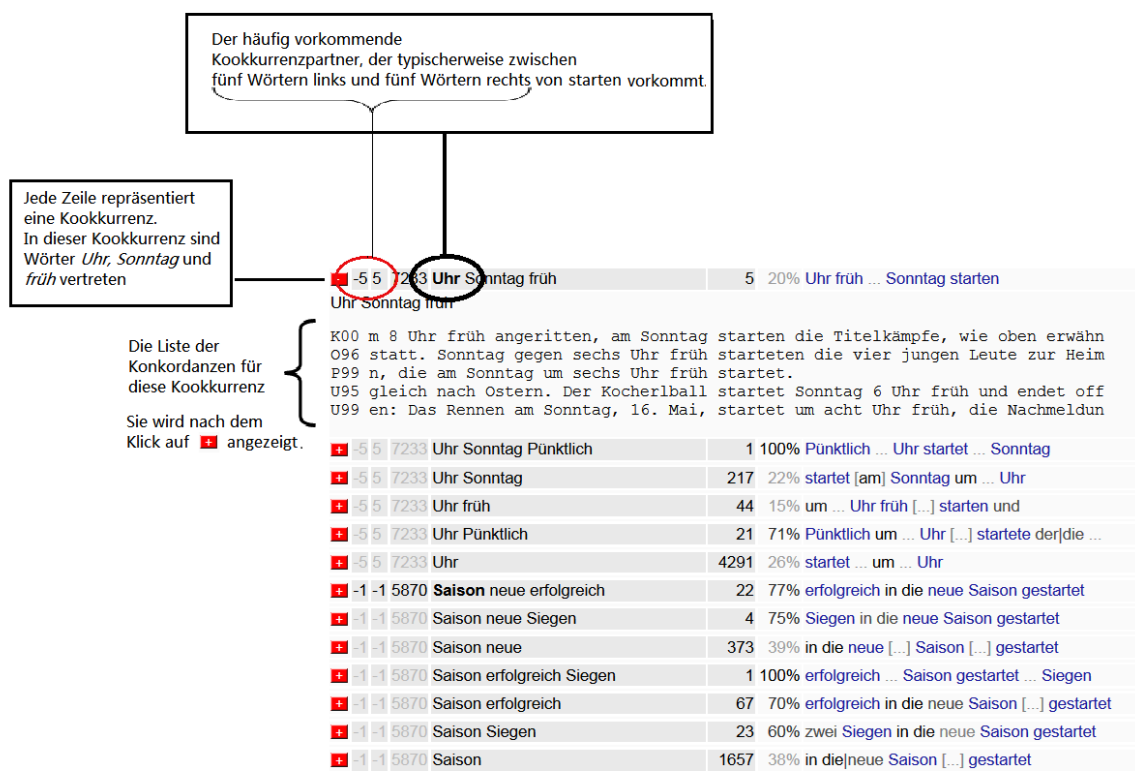


Abb. 2: Ausschnitt aus der Kookkurrenzanalyse. Detailliert beschrieben unter <http://corpora.ids-mannheim.de/ccdb/>. Meine Grafik, H. S.

Das DeReKo beinhaltet eine Menge von Belegen eines analysierten Objekts, die es in authentischen Texten präsentieren. Die gesuchten Objekte sind quer durch das ganze

Korpus verstreut. Der Algorithmus der CCDB ermöglicht das Korpus zu recherchieren und statistisch auszuwerten. Dann kann festgelegt werden, was für den Gebrauch des analysierten Lexems prototypisch ist (vgl. Vachková, Belica 2009: S. 226). Außerdem ist auch der Grad der Überlappung von Kookkurrenzprofilen verschiedener Wörter mithilfe der CCDB zu beobachten (CNS, s. u.). Das hilft bei der Kontrastierung zweier Wörter, ihrer Kontexte und Sprachverwendung. Die Grade der Überlappung sind unterschiedlich hoch und die CCDB ist in der Lage sie zu veranschaulichen. Dank dieser Veranschaulichung ist schnell festzustellen, inwieweit sich die KP der kontrastierten Wörter ähneln oder unterscheiden. Dazu verhelfen gleichfalls die Merkmalskarten (s. u.). Sie dienen zur schnellen Orientierung, die am Anfang jeder Untersuchung erforderlich ist.

Die Höhe eines Überlappungsgrades wird nach dem gemeinsamen Auftreten von Wörtern in den gleichen Verbindungen gemessen. Es wird berücksichtigt, wie oft die Wörter in welchen Kontexten miteinander vorkommen. Im Prinzip lässt sich jede lexikalische Einheit durch ein eigenes KP charakterisieren. Ein KP jedes Wortes besteht aus einer Menge von Lexemen, die für das Wort prototypisch sind (Vachková, Belica 2009: S. 224).

Das Kookkurrenzprofil eines beliebigen Wortes entsteht folgendermaßen (vereinfacht erklärt, vgl. Perkuhn et al. 2012: S. 116ff): Das COSMAS-System (d. h. das System, das eine Recherche im DeReKo und anderen Korpora ermöglicht) durchsucht die sich im DeReKo befindenden Texte, in denen sich das ausgewählte Wort (sog. Bezugswort) befindet. Das System registriert Positionen und benachbarte Wörter, in deren lokalem Kontext das untersuchte Wort vorkommt. Dann wird vom System festgelegt, „wie viele Wörter davor und danach für die Fragestellung berücksichtigt werden sollen“ (ibidem). Das Verfahren wertet die Kontexte mit allen Wortformen aus. Dabei ist bekannt, wie oft diese Wortformen im Korpus vorkommen und welche ihre genaue Anzahl im Gesamtkorpus und die Anzahl in der Gesamtheit der definierten Trefferkontexte ist (ibidem). Das KP wird mithilfe des Systems strukturiert, hierarchisiert und danach visualisiert (s. u.).

3.2.4. Merkmalskarte

Die KPe können in der CCDB übersichtlich strukturiert und in sog. Merkmalskarten (weiter nur MK) mit Farben visualisiert werden (s. Abb. 3). Das System generiert eine

„Tabelle“ (MK), welche die KPe gruppiert und veranschaulicht. Es ermöglicht einen (meistens) schnellen und detaillierten Einblick in die Bedeutungsstruktur eines Lexems. Die MK dient zur Untersuchung von Lexemen, ihrer praktischen Anwendung und Beziehung zu anderen Lexemen. Mittels der MK lässt sich feststellen, in welchen Kontexten die Lexeme am häufigsten erscheinen. Bei der Erforschung von Lexemen sollte die lexikalische Umgebung (der lokale Kontext) in Betracht gezogen werden. Das bedeutet, wenn ein Lexem (seine Bedeutung, stilistische Prägung usw.) untersucht wird, dann sollten auch die mit ihm vorkommenden Wörter berücksichtigt werden. Die MKen sind bei solchen Untersuchungen sehr hilfreich.

Die MK bilden 25 Quadrate (fünf mal fünf). Manchmal ist das System aufgrund der Umstände nicht in der Lage, gleichzeitig alle Quadrate anzulegen. In solchen Fällen werden weniger Quadrate gebildet und so können leere Quadrate entstehen (vgl. Perkuhn et al. 2012: S. 135).

Perkuhn fügt hinzu:

„Intuitiv kann man dies so deuten, dass die um ein derartiges Quadrat gelegenen Cluster sich zu wenig mögen, sich zu unähnlich sind, als dass sie es hinnehmen könnten, nebeneinander platziert zu werden.“ (ibidem).

Auch Vachková (2007a: S. 47) weist darauf hin. Wenn eine MK Leerquadrate beinhaltet, kann das bedeuten, dass sie zwei (oder mehr) Lesarten im Schema signalisieren. Sie entstehen deshalb sehr oft bei homonymischen Lexemen.

Die Anordnung der Quadrate ist zu einem gewissen Grad zufällig. Jedes Quadrat enthält eine Lexemgruppe (sog. Cluster), die einen gemeinsamen lokalen Kontext (KP) vertreten und die auch dem lokalen Kontext des Bezugsworts ähneln. Sie sind (obwohl nicht auf Anhieb) gewissermaßen miteinander semantisch verwandt (s. o.). Dabei spielen die Wortarten und -formen keine Rolle. Trotzdem gibt es einige Beschränkungen. Zum Beispiel treten Pronomen, die zu Synsemantika zählen, nur selten auf, da sie semantisch verblasst sind und keine lexikalische Bedeutung aufweisen. Dagegen kommen Interjektionen oder onomatopoetische Verben häufig vor, denn sie haben eine offenbare Bedeutung (vgl. Vachková, Belica 2009: S. 229). Die syntaktischen Strukturen sind unterdrückt und nicht immer ablesbar (vgl. Vachková 2008: S. 197), aber zuweilen können sie in einzelnen Quadraten widerspiegelt werden (s. MK 1, E3).

Aus jedem Quadrat lässt sich ein globaler situativer Kontext entschlüsseln. Die Lexeme hängen zusammen (auffällig oder getarnt) und bilden den Themenbereich des jeweiligen Quadrats. Es ist nicht immer einfach, einen globalen Kontext auf dem ersten Blick zu erkennen. Zunächst muss der lokale Kontext analysiert werden, erst dann ist der globale Kontext zu bestimmen. Während der Analyse wird das Prinzip der Assoziation angewendet. Beim Blick auf ein Quadrat sollen die Ideen „frei strömen“ und man darf sich nicht lediglich auf einzelne Wörter konzentrieren, sondern es muss die ganze Gruppe als ein Komplex verstanden werden. Um eine MK bzw. ein Quadrat erfolgreich zu analysieren, muss der Forscher über gewisse sozio-kulturelle Kenntnisse verfügen. Auch linguistische Kenntnisse können bei der Analyse bedeutend helfen. Die Interpretation hängt jedoch deutlich vom individuellen Mentallexikon ab. Zugleich gilt aber, dass Personen aus demselben (sozio)kulturellen Raum ähnliche Mentallexikone haben, sodass eine intersubjektive Interpretation der Merkmalskarten möglich ist. Vor allem die Nichtmuttersprachler können aber bei der Analyse Schwierigkeiten haben, aber auch für Muttersprachler ist die Analyse nicht immer problemlos. (vgl. Vachková, Belica 2009: S. 228ff).

Dieses Quadrat enthält eine **Lexemgruppe** von Wörtern der Wochentage.

Diese Wörter besitzen ein KP, das ihnen aus großem Teil gemeinsam ist, d. h. sie kommen in ähnlichen lokalen Kontexten vor, die überdies dem KP von "starten" ähneln.

Der **Cluster** (d. h. eine Gruppe von Wörtern, die miteinander zusammenhängen), bezieht sich hier auf Zeitangaben, auf die Zeit, in der etw. "startet".

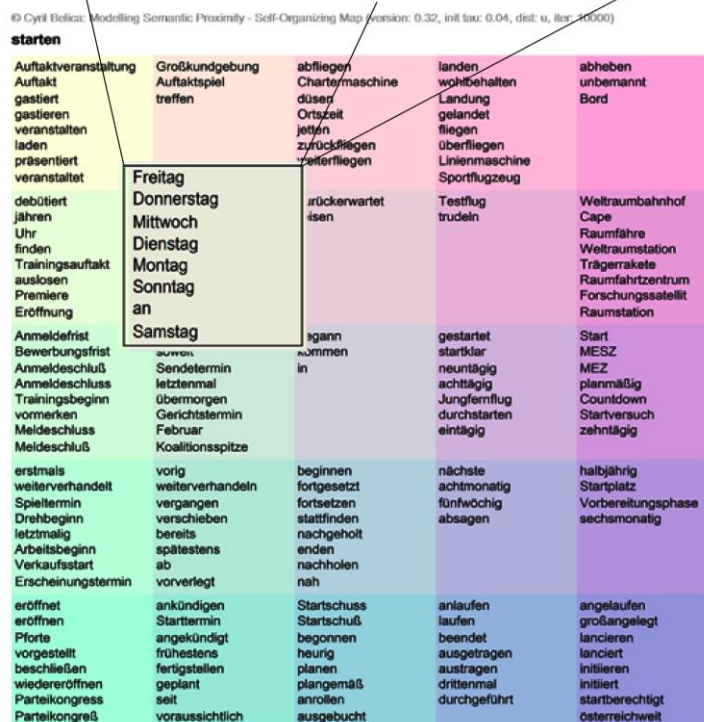


Abb. 3: Selbstorganisierende lexikalische Merkmalskarte (SOM) für das Lexem **starten**.

Daraus ergibt sich, dass die MKen semiotisch interpretierbar sind, denn sie illustrieren lexikalische Beziehungen zwischen Lexemen. Die lexikalischen Einheiten werden im *Parole*-Bereich (s. o.) untersucht, d. h. in der Realwelt, in einer konkreten Situation usw. (vgl. ibidem: S. 223). Die MKen spiegeln also die Verwendung eines Lexems in Realität wider. Es wird gezeigt, in welchem Diskurs das Lexem eingebunden ist. Das spielt für seinen Gebrauch in der Praxis eine wichtige Rolle (besonders für einen Nichtmuttersprachler), denn in Wörterbüchern sind die Wörter oft aus dem Kontext gerissen.

Die CCDB stellt zwei Typen von Merkmalskarten zur Verfügung:

- SOM – Self-Organizing Map
- CNS – Contrasting Near-Synonyms

In nächsten Absätzen wird beschrieben, wie sich die konkreten Merkmalskarten interpretieren lassen.

3.2.4.1. Self-Organizing Map (SOM-Merkmalskarte)

Die SOM-Merkmalskarte (s. o., Abb. 3) ist ein strukturiertes und visualisiertes Kookkurrenzprofil eines Lexems. Die Anordnung der Farbtöne ist festgelegt und sie soll die Ähnlichkeit der Profilquadrate auch als Ähnlichkeit zwischen den Farben widerspiegeln (Perkuhn et al. 2012: S. 134.). (SOM wurde anhand des Kohonennetzes entwickelt, s. dazu Kohonen 1995.)

Es ist ein bestimmtes Verfahren bei der SOM-Interpretation zu empfehlen, das ihre Analyse erleichtert (vgl. Vachková, Belica 2009: S. 228ff und Vachková et al. 2007: S. 19). Mit der Analyse der Merkmalskarte sollte man in einer beliebigen Ecke der MK anfangen. Der Nutzer versucht aufgrund seiner Assoziationen den „verborgenen“ Sinn zu enthüllen (s. o.) und zwar nach semantischen oder syntaktischen Kriterien. Die Bestimmung des thematischen Bereichs wird meistens nach Zugehörigkeit zu einem Wortfeld festgelegt. Bei der Interpretation ist es nötig, außerhalb der linguistischen Kategorien nachzudenken und eigene bisherige Erkenntnisse zu vermeiden, denn sie verhindern freies Denken ohne Einschränkungen. Wenn das erste Quadrat entschlüsselt wird, soll es durch die Untersuchung der benachbarten Quadrate erweitert werden. Es

Dieses Verfahren wird in alle Richtungen und von allen Ecken wiederholt. Die endgültige Auswertung wird hinsichtlich einzelner Quadrate durchgeführt (oft mit Rücksicht auf lexikografische Kriterien), wobei nicht immer alle Quadrate analysiert werden müssen, um ein befriedigendes Ergebnis der Untersuchung zu erreichen.

Die Methode CNS dient dazu, die miteinander verwandten KP zweier beliebiger Wörter gegenseitig zu kontrastieren. Es wird eine Merkmalskarte generiert, in der quasi zwei SOM-Merkmalskarten nebeneinander stehen (bei Wörtern, deren KP zu unterschiedlich sind, kann das System jedoch keine CNS-Merkmalskarte erstellen). Der Unterschied zur SOM besteht u. a. darin, dass die Farben eine andere Funktion haben, sie visualisieren nämlich den Kontrast zwischen den KPen der zwei verglichenen Wörter. In der CNS-Merkmalskarte gibt es darum verschiedene Farbtöne zwischen der roten und gelben Farbe – für ein untersuchtes Wort jeweils eine bestimmte Farbe (entweder rot oder gelb). Die Intensität der Farben illustriert das Maß der Überlappung von KPen (Abb. 4).

© Cyril Bolcar: Modeling Semantic Proximity – Contrasting Near-Synonyms (version 0.2.1, last Dec. 6/8, 2017, X, last: 10/09/20)

The diagram illustrates the semantic proximity of 'beginnen' and 'anfangen' through a network of related verbs. The central grid is organized as follows:

laden	nah	schreien	ficken	fangen
Vormittag	Ende	plänen	meckern	aufwachen
finden	Oktober	heulen	kecken	ankommen
Abend	Badesaison	saufen	lachen	los
Morgenstunde	Spätherbst	krähen	atmen	lose
Abschlussveranstaltung	September	flüstern	schießen	durchstarten
Abschlussveranstaltung	April	schneien	schlafen	aufstehen
Uhr	Februar	jammern	Schieß	aufgewacht
Anmeldefrist	voraussichtlich	losgehen	hinsetzen	angefangen
treten	frühestens	loslegen	quatschen	gewöhnen
Bewerbungsrlist	fortigstellen		reden	Fangen
Freitag	Baubeginn	kommen	immerzu	beginne
sowie	spätestens	dämmern	schwätzen	Zähneputzen
Mittwoch	ab		fernsehen	radfahren
Donnerstag	seit		dasitzen	
Dienstag	Spätestlich		zusammensitzen	
Startschuß	begonnen		angewöhnen	lernen
starten	abgeschlossen		Aufhören	Theaterspiell
daraufolgend	anlaufen		zuhören	verlernt
stattfinden	abschließen		aufhören	verlernen
nachgehoit	Hochtour		abgewöhnen	Spaß
eröffnet	planmäßig		anstatt	Tänzen
eröffnen	angelaufen		Angewöhnheit	töpfen
			angewöhnt	Fußballspielen
begann				
enden				
dauern				
beendet				
nächste				
fortsetzen				
fortgesetzt				
Beginn				
beenden				
abbrechen				
abgebrochen				
Beginnen				
mehrwöchig				
fünfwoöchig				
viermonatig				
fünfmonatig				

Arrows point from the central grid to two peripheral boxes:

- ANFANGEN** (orange box): fangen, aufwachen, ankommen, los, lose, durchstarten, aufstehen, aufgewacht, angefangen, gewöhnen, Fangen, beginne, Zähneputzen, radfahren.
- BEGINNEN** (red box): laden, nah, Ende, Oktober, Badesaison, Spätherbst, September, April, Februar, voraussichtlich, frühestens, fortigstellen, Baubeginn, spätestens, ab, seit, Spätestlich, begonnen, abgeschlossen, anlaufen, abschließen, Hochtour, planmäßig, angelaufen, beenden, abbrechen, abgebrochen, Beginnen, mehrwöchig, fünfwoöchig, viermonatig, fünfmonatig.

27

Der Vergleich von KPen beruht auf der Veranschaulichung von Überlappungen (oder Unvereinbarkeiten) zweier KPe und die Merkmalskarte ermittelt entweder Ähnlichkeiten oder Differenzen in der Verwendung der zwei Wörter. Anhand der CNS werden empirische Einsichten in die Zusammenhänge zwischen den verglichenen Wörtern hinsichtlich der lexikalisch-semanticen Beziehung erforscht. Das hilft wieder – aufgrund der globalen Kontexte (bzw. lokalen Kontexte) – feine Bedeutungsunterschiede zu entschlüsseln.

Wie es sich aus diesen Beschreibungen und schon aus der eigentlichen Bezeichnung der Methode (*Contrasting Near-Synonyms*) voraussetzen lässt, wird die CNS-Merkmalskarte besonders zum Vergleichen zweier Synonyme benutzt, vor allem wenn Unsicherheiten in ihrer Verwendung bestehen. Trotzdem müssen nicht nur Synonyme verglichen werden.

„Die Kookkurrenznähe impliziert nicht nur ‚semantische Nähe allein‘, sie betrifft u. a. die Paradigmatik. Die Konstellation der KP innerhalb eines Clusters kann darauf hindeuten, dass in bestimmten Diskursen zwar unterschiedliche lexikalische Relationen vorkommen, manche jedoch als besonders diskurstypisch erscheinen.“ (Vachková 2008: S. 200)

Laut dieser Aussage muss es sich nicht unbedingt um Synonyme handeln, wenn die Lexeme in gleichen Kontexten erscheinen. Genauso bezieht sich diese Theorie auf Antonyme, bzw. Hyponyme, Hyperonyme und Kohyponyme, d. h. auf Beziehungen zwischen Wörtern, die auf einer paradigmatischen Ebene wirken.

Damit hängt ebenso der stilistische Wert des Lexems zusammen. Die stilistische Funktion ist aus der MK nicht immer detailliert ablesbar. Auf der anderen Seite muss man auf die Tatsache Bezug nehmen, dass *„ein Lexem mit der bestimmten stilistischen Funktion Lexeme mit gleicher stilistischer Funktion anzieht“* (Marková 2012b: S.63). Daraus kann eventuell eine stilistische Ebene festgestellt werden (wie z. B. umgangssprachliche oder derbe Ausdrücke u. Ä.).

4. Empirischer Teil – Analyse der synonymischen Reihen

Im folgenden Teil der Bachelorarbeit werden die Analysen zweier synonymischen Reihen mithilfe der Merkmalskarten durchgeführt:

- anfangen – beginnen – starten
- herstellen – erzeugen – produzieren

Diese Verben habe ich deshalb ausgewählt, weil sich besonders die Nichtmuttersprachler mit Unklarheiten bei der Verwendung häufig auseinandersetzen müssen. Oft werden diese Verben von Nichtmuttersprachlern nämlich als absolute Synonyme verstanden. Vor allem in den zweisprachigen Wörterbüchern und sogar in vielen Synonymwörterbüchern fehlen meistens ausführlichere Angaben und konkrete Beispiele, die solche Nuancen zwischen den oben angeführten Verben klären.

4.1. Analyse der Synonymreihe: ANFANGEN – BEGINNEN – STARTEN

ANFANGEN <st. V.; hat> [mhd. an[e]vāhen, ahd. anafāhan, urspr. = anfassen, in die Hand nehmen]	BEGINNEN <st. V.; hat> [mhd. beginnen, ahd. beginnan, zu einem nur in Zus. erhaltenen germ. Verb, vgl. got. duginnan = beginnen]	STARTEN <sw. V.> [engl. to start = fort-, losgehen, -fahren]
<p>1. a) <i>etw. in Angriff nehmen, mit etw. beginnen:</i> eine Arbeit, einen Brief, eine Freundschaft, ein Gespräch, ein neues Leben a.; sie fing wieder an zu paddeln/fing wieder zu paddeln an; gleich werden die Sirenen zu heulen a./a. zu heulen; du hast angefangen (ugs.; <i>hast den Streit o. Ä. begonnen</i>); ein Verhältnis, etwas mit jmdm. a. (ugs.; <i>eine Liebesbeziehung mit jmdm. beginnen</i>);</p> <p>b) <i>eine Ausbildung, eine berufliche Arbeit beginnen:</i> am 1. März können Sie a.; [ganz] von vorn[e], von klein auf a. (<i>mit dem untersten, am schlechtesten bezahlten Posten beginnen</i>); sie hat als Handelsvertreterin angefangen;</p> <p>c) <i>zu reden beginnen:</i> »Liebe Freundinnen und Freunde«, fing er an; sie fing mit diesen Worten an;</p> <p>d) (ugs.) <i>ein bestimmtes Thema anschneiden:</i> er fing immer wieder von Politik an; nun fang du auch noch damit/davon an!</p> <p>2. a) <i>zu etw. gebrauchen, anstellen:</i> nichts, etwas mit sich, mit seiner Freizeit anzufangen wissen; mit ihm ist heute nichts anzufangen (<i>er ist heute nicht in Form, nicht ansprechbar</i>);</p> <p>b) <i>machen, tun:</i> was können, sollen wir nachher a.?; eine Sache richtig, verkehrt a.</p> <p>3. [mit etw.] <i>einsetzen, beginnen, seinen Anfang nehmen:</i> hier fängt das Sperrgebiet an; der Unterricht fing um halb neun an; das Wort fängt mit p an; (iron.): das fängt ja gut, schön, heiter, nett an.</p>	<p>1. a) <i>mit etw. einsetzen, einen Anfang machen; anfangen:</i> eine Arbeit b.; er hat [ganz unten] als Laufbursche begonnen; sie beginnt wieder mit dem Training;</p> <p>b) <i>auf bestimmte Weise tun, unternehmen, anstellen:</i> wir müssen die Sache anders b.</p> <p>2. <i>seinen Anfang haben; zu einer bestimmten Zeit, an einem bestimmten Ort, auf bestimmte Weise anfangen:</i> die Vorstellung beginnt um 20 Uhr; Namen, die mit dem Buchstaben B beginnen; dort hinten beginnt die Schweiz; es begann zu schneien; im beginnenden 21. Jahrhundert.</p>	<p>1. a) <hat> <i>einen Wettkampf (Wettlauf, -rennen, -schwimmen o. Ä.) beginnen lassen:</i> das Autorennen s.;</p> <p>b) <ist> (<i>auf ein akustisches, auch optisches Signal hin</i>) <i>einen Wettkampf (Wettlauf, -rennen, -schwimmen o. Ä.) beginnen:</i> zur letzten Etappe s.; Ü der Außen war schneller gestartet (<i>schneller losgelaufen</i>) als der Verteidiger;</p> <p>c) <ist> <i>an einem Wettkampf aktiv teilnehmen:</i> bei einem Wettkampf s.</p> <p>2. a) <hat> <i>bewirken, dass etw. auf ein Ziel hin in Bewegung gesetzt wird:</i> eine Rakete s.;</p> <p>b) <ist> <i>sich (irgendwohin) in Bewegung setzen:</i> das Flugzeug ist pünktlich gestartet.</p> <p>3. a) <hat> (<i>durch Betätigung einer Taste, des Anlassers o. Ä.</i>) <i>in Gang setzen, in Betrieb nehmen:</i> den Motor, das Auto s.;</p> <p>b) (<i>durch Betätigung einer Taste, des Anlassers o. Ä.</i>) <i>in Gang gesetzt werden, in Betrieb genommen werden:</i> der Computer startet.</p> <p>4. a) <hat> (<i>eine Unternehmung, ein Vorhaben o. Ä.</i>) <i>beginnen lassen:</i> eine große Aktion s.;</p> <p>b) <ist> <i>aufbrechen, um eine Unternehmung, ein Vorhaben o. Ä. durchzuführen:</i> sie sind gestern [in den Urlaub, zu einer Expedition] gestartet;</p> <p>c) <ist, selten: hat> <i>anlaufen, seinen Anfang nehmen, beginnen:</i> die Tournee startet in Hamburg.</p>

Tabelle 1: Angaben in DUW

4.1.1. Analyse *anfangen* – *beginnen*

Die Verben *anfangen* und *beginnen* sind nahe Synonyme, wobei der Unterschied zwischen ihnen nicht leicht festzustellen ist und man an ihre Verwendung zweifeln kann. In Linke et al. (2004: S. 170) ist erwähnt, dass diese Verben in verschiedenen Aspekten ganz synonymisch sind: regional, fachsprachlich/gemeinsprachlich, soziolektal, stilistisch (d. h. in allen Textsorten) und in der Kollokation. Diese Verben waren ursprünglich Regionalismen – im Süden wurde *anfangen*, im Norden *beginnen* benutzt. In der Gegenwart sind sie „im ganzen deutschen Sprachraum verbreitet, ohne dass zwischen ihnen eine Bedeutungsdivergenz festzustellen wäre“ (Linke et al. 2004: S. 170). Laut solcher Aussage scheinen sie rein synonymisch zu sein, d. h. bedeutungsgleich, was eben für regionale Synonymie typisch ist.

Heusinger (2004: S. 178) behauptet, dass die bedeutungsidentischen Lexeme frei substituierbar sein müssen. Als Beispiel führt Heusinger gerade das Synonympaar *anfangen* und *beginnen*. Seine Frage ist, ob sich die Verben in jedem Kontext austauschen lassen.

„Bei Sememidentität wird die Norm nicht verletzt (den Brief angefangen/ den Brief begonnen haben), doch handelt es sich nur um Sememaffinität, dann ist die Möglichkeit des Austausches nicht mehr uneingeschränkt gegeben (z. B. im Sinne von ‚begreifen‘: ‚mit der Aufgabe/ der Frage nichts anfangen können‘)“. (ibidem)

Die freie Austauschbarkeit kann entweder von Konventionen oder vom Verwendungssus eingeschränkt sein. Es werden kommunikative und soziale Bedingungen festgelegt (wie z. B. alltägliche Kommunikation, wissenschaftliche, politische oder offizielle Sprache und Umgebungssituation, d. h. außersprachliche Wirklichkeit). Heusingers Aussage unterscheidet sich darum völlig von der oben erwähnten These von Linke et al.

Außerdem führt Heusinger (2004: S. 179) Beispiele an, in denen bestimmte Restriktionen zu beobachten sind, wobei – wie er schreibt – die semantischen Argumente aus synchronischer Sicht zu den Einschränkungen nicht bekannt sind.

- a) *der Anfang wurde gemacht*
- b) ** der Beginn wurde gemacht*
- c) *mit dem Bau wurde begonnen*
- d) ** mit dem Bau wurde angefangen*

Weiter gibt er noch ein anderes Beispiel an:

Die Ferien/ der Winter/ die Sportveranstaltung/ die Saison (haben) hat begonnen/ (haben) hat angefangen. (ibidem: S. 183)

Für dieses Beispiel gibt es nach Heusinger keine kollokativen Beschränkungen. Die CNS-Merkmalkarte zeigt aber deutlich, dass ausschließlich das Verb *beginnen* in der Kookkurrenz von Lexemen der Zeitabschnitte erscheint, d. h. es wird in den realisierten Aussagen üblicherweise anders verwendet, als er behauptet (s. u.).

Heusinger führt noch ein bedeutendes Beispiel an, das mit den Angaben aus der CNS-Merkmalkarte übereinstimmt.

der Bau beginnt vs. * *der Bau fängt an* (ibidem)

Hier handelt es sich um die Bedeutung im Sinne „ein Vorhaben wird realisiert“, wie es Heusinger beschreibt und wie DUW anführt, d. h. ein geplantes Ziel wird verwirklicht. In diesem Fall können die Verben nicht substituiert werden. (Konkrete Beispiele siehe die CNS-Analyse, MK1)

Analyse der Merkmalskarte

© Cyril Belica: Modelling Semantic Proximity – Contrasting Near-Synonyms (version: 0.21, init tau: 0.4, dist x, iter: 10000)

anfangen	beginnen				
laden			schreien	ficken	fangen
Vormittag			plärren	meckern	aufwachen
finden			heulen	kotzen	ankommen
Abend			saufen	lachen	los
Morgenstunde			krähen	atmen	lose
Abschlussveranstaltung			flüstern	scheißen	durchstarten
Abschlußveranstaltung			schneien	schlafen	aufstehen
Uhr			jammern	Scheiß	aufgewacht
Anmeldefrist	nah		losgehen	hinsetzen	angefangen
treten	Ende		loslegen	quatschen	gewöhnen
Bewerbungsfrist	Oktober			reden	Fangen
Freitag	Badesaison			immerzu	begänne
soweit	Spätherbst			schwätzen	Zähneputzen
Mittwoch	September			fernsehen	radfahren
Donnerstag	April			dasitzen	
Dienstag	Februar			zusammensitzen	
Startschuß	voraussichtlich		kommen	angewöhnen	lernen
Startschuss	frühestens		dämmern	Aufhören	Theaterspiel
starten	fertigstellen			zuhören	verlernt
darauffolgend	Baubeginn			aufhören	verlernen
stattfinden	spätestens			abgewöhnen	Spaß
nachgeholt	ab			anstatt	Tanzen
eröffnet	seit			Angewohnheit	töpfern
eröffnen	Spatenstich			angewöhnt	Fußballspielen
begann	begonnen			gerne	mitkriegen
enden	abgeschlossen			geschweige	kapiert
dauern	anlaufen			selber	zumute
beendet	abschließen			dauernd	merken
nächste	Hochtour			gern	kapieren
fortsetzen	Hochtouren			können	mitbekommen
fortgesetzt	plangemäß			genausogut	Ahnung
Beginn	angelaufen			Leute	begriffen
beenden			worüber	ja	rauskommen
abbrechen			was	tun	rausgehen
abgebrochen			woran	gar	sowieso
Beginnen			wen	irgendwas	dran
mehrwöchig			wovon	schiefgehen	hinschmeißen
fünfwöchig			worum	eingefallen	hierherkommen
viermonatig			wem	irgend	nachher
fünfmonatig			wer	jemand	kriegen

MK1

Schon auf den ersten Blick auf die CNS-Merkmalskarte ist offensichtlich, dass diese zwei Verben gar nicht rein synonymisch sein können. Es werden mithilfe der roten und gelben Farbe ganz deutlich zwei unterschiedliche Kookkurrenzprofile gegeneinander gestellt. Orange Farbe, die eine Überlappung signalisiert, ist hier nur mit vier Quadraten repräsentiert. Daraus lässt sich schließen, dass die Verben nicht immer willkürlich austauschbar sind und dass sie in jeweils unterschiedlichen Kontexten auftauchen.

Auf der linken Seite ist das KP des Verbs **beginnen** zu beobachten. *Beginnen* wird in dem Kontext einer Veranstaltung benutzt (wie aus Quadraten A1, B1 und C1 ersichtlich ist),

die zu einer bestimmten Zeit stattfindet (vgl. Quadrate A1 und B1, welche die Lexeme der Zeitangaben beinhalten). Bei C1 könnte es sich um einen Kampf, bzw. sportlichen Wettkampf handeln, aber auch andere Typen der Veranstaltungen. Quadrate D1, E1 signalisieren einen Vorgang, wobei im E1 solche Wörter präsent sind, die gleichfalls mit Zeitangaben zusammenhängen, und zwar mit der Bedeutung von Dauer (*mehrwöchig*, *viermonatig* u. a.). Die nächste vertikale Achse (2) zeigt wieder einen Kontext der Zeitangaben (sowohl konkrete Monate als auch Zeitabschnitte wie *Spätherbst* oder *Badesaison*). Das steht im Einklang mit den Angaben in DUW begründet, in dem *beginnen* mit einem Anfang verknüpft ist, der „zu einer bestimmten Zeit, an einem bestimmten Ort, auf bestimmte Weise“ verläuft (s. Tabelle 1). Auf der anderen Seite entspricht die Merkmalskarte nicht den Beispielen von Heusinger, in denen er Zeitangaben nicht nur mit *beginnen*, sondern auch mit *anfangen* verbindet (s. o.). Im Quadrat C2 ist der Kontext der Renovierung/Bauarbeit beinhaltet, was Heusinger als ein Beispiel angeführt hat (s. o.), in dem *anfangen* in diesem Kontext ausgeschlossen ist. Der globale Kontext Bauarbeit und auch andere Typen der Arbeiten beziehen sich ebenso auf das Quadrat D2. Zur Interpretierung dieses Quadrats trägt das Substantiv *Hochtour(en)* bedeutsam bei. In DUW ist folgende Redewendung zu finden: *auf Hochtouren laufen/arbeiten* (d. h. „mit der größten Leistungsfähigkeit laufen/arbeiten; unter großer Hektik u. unter Aufbringung aller Kraftreserven vorstattengehen“). Daraus folgt, dass es in D2 (auch aufgrund anderer Wörter des Quadrats) um den Kontext der Arbeit oder Vorbereitung geht.

Die rechte Seite der Merkmalskarte zeigt vor allem KPe solcher Wörter, die dem KP von *anfangen* ähneln. *Anfangen* ist offensichtlich mit menschlichen Geräuschen (A3 – *schreien*, *flüstern* u. a.), Ausdrücken (B4 – *reden*, *quatschen* u. a.) und Tätigkeiten (bzw. gesellschaftlichen Ereignissen) verbunden (B4 – *fernsehen*, C5 – *lernen*, *Tanzen*, *Fußballspielen* u. a.). Obgleich die Verben wie *krähen*, *heulen* (A3) oder *meckern* (A4) Tiergeräusche evozieren, können sie auch menschliche Äußerungen ausdrücken (und zwar im metaphorischen Sinne). Das wird mit den folgenden Korpusbelegen aus der CCDB untermauert:

Marie sah aus, als würde sie gleich anfangen zu heulen.

[...]Ein Junge kräht dazwischen.[...]

[...]„Die Leute haben immer etwas zu meckern“[...]

Auf der anderen Seite ist in der CCDB ein Beispiel zu finden, das der These widerspricht:

[...]dann *beginnt* man zu krähen[...]

Allerdings muss daran erinnert werden, dass man anhand geringer Anzahl von sprachlichen Belegen im Korpus keine eindeutigen Schlussfolgerungen ziehen darf (s. o.).

Zu weiteren Kontexten, die auf Menschen bezogen sind, gehören Quadrate D5 und teilweise C4. Dem Quadrat D5 lässt sich der Kontext des Verständnisses entnehmen (*mitkriegen, kapieren, begriffen* usw.). Quadrat C4 mit Wörtern wie *angewöhnt, aufhören* u. a. stellt wahrscheinlich Sucht/Abhängigkeit dar. Ganz klar ersichtlich ist ein syntaktisches Merkmal, das dem Verb *anfangen* zugehört (E3 – *was, woran, wen, wer* u. a.). Dies könnte man so verstehen, dass das Verb *anfangen* vor dem Verb *beginnen* in W-Fragen bevorzugt wird. Jedoch lässt sich diese Vermutung nicht klar beweisen. Durch die Internetsuchmaschine werden es sogar mehrere Ergebnisse von W-Fragen mit *beginnen* angezeigt als mit *anfangen*. Was aber die Wörter in E3 und das Wort *anfangen* gemeinsam haben, ist das häufig vorkommende Kookkurrenzpartner *wissen* (in verschiedenen Formen). Vielleicht deutet dieses Quadrat auf indirekte Fragen oder Nebensätze hin, wie in den folgenden Beispielen aus der CCDB ersichtlich ist:

Warum ich anfang, Geschichten zu schreiben, weiß ich [...]

Wer den Streit anfang, wußte nachher niemand mehr. [...]

[...] und wußte nicht, was er anfangen sollte. [...]

In der Mitte und in der fünften vertikalen Achse befinden sich Quadrate, die eine **Überlappung** von KPen der Verben darstellen. Es handelt sich um den Kontext des angehenden Tages bzw. der *Nacht* (C3 – *kommen, dämmern*). Verben in Quadrat B3 (*losgehen* und *loslegen*) sind umgangssprachliche Synonyme zu *anfangen* und *beginnen*.

Beispiele aus der CCDB:

[...] Dann soll die Party richtig losgehen. [...]

Als es losging, war es elf Uhr morgens.

Erst um sechs Uhr wird er richtig loslegen. [...]

Zusammenfassung:

Obwohl die Verben *anfangen* und *beginnen* sehr nahe Synonyme und nach Linke et al. (2004) sogar bedeutungsidentisch sind, haben sie unterschiedliche Kookkurrenzprofile und deshalb auch Verwendungsweisen. Während *beginnen* v. a. im Zusammenhang mit Zeitangaben und eher offiziellen Veranstaltungen verwendet wird, gehört *anfangen* zu Kontexten der menschlichen Äußerungen und Handlungen. Dabei ist wichtig zu bemerken, dass es dem Verb *anfangen* auffällig viele Kookkurrenzprofile der umgangssprachlichen – sogar vulgären – Ausdrücke ähneln (z. B. A4 – *kotzen*, *scheißen*, B4 – *quatschen*, *schwätzen*, D5 – *mitkriegen*, *kapieren*, E4 – *irgendwas*, *schiefgehen*, E5 – *rausgehen*, *rauskommen*, *hinschmeißen*, *hierherkommen* u. a.).

4.1.2. Analyse *beginnen*-*starten*

© Cyril Belica: Modelling Semantic Proximity - Contrasting Near-Synonyms (version: 0.21, init tau: 0.4, dist: x, iter: 10000)

starten	beginnen			
Weltraumbahnhof	abheben	abfliegen	reisen	startberechtigt
Cape	landen	Chartermaschine	treffen	debütiert
Raumfähre	wohlbehalten	fliegen	in	präsentiert
Weltraumstation	Landung	überfliegen		vorgestellt
Trägerrakete	gelandet	durchstarten		ausgetragen
Raumfahrtzentrum	Bord	Linienmaschine		letztmalig
Forschungssatellit	trudeln	düsen		Großdemonstration
Raumstation	Flug	Sportflugzeug		Verkaufsstart
Countdown	Start	zurückerkwartet	fortgesetzt	
Startplatz	MESZ	startklar	fortsetzen	
zehntägig	MEZ		treten	
	planmäßig		stattfinden	
	Startversuch		enden	
	achttagig		weiterverhandeln	
	Testflug		beschließen	
	gestartet		terminieren	
großangelegt	Startschuß	vorig	eröffnet	Freitag
lancieren	Startschuss	bereits	Anmeldefrist	Donnerstag
lanciert	anlaufen	spätestens	eröffnen	Mittwoch
initiiieren	begonnen	vergangen	Bewerbungsfrist	Dienstag
initiiert	nächste	plangemäß	soweit	Montag
angelaufen	Auftakt	nachgeholt	erstmal	laden
halbjährig	laufen	ab	Sendetermin	Sonntag
österreichweit	heurig	nachholen	wiedereröffnen	Uhr
Beginn	angefangen	abschließen	zusammentreten	Weltkindertag
eingeläutet	anfangen	Hochtour	ansetzen	
Beginnen	fangen	Hochturen	verkündet	
angebrochen	losgehen	fertig	Friedenspflicht	
mitten	begann	hinziehen	bis	
beenden	aufhören	anstehend	Erstnotiz	
einstimmen	anbrechen	anrollen	Arbeitsbeginn	
einläuten	kommen	loslegen	Neumond	
abbrechen	dauern	Baubeginn	Oktober	Vormittag
mehrwöchig	abgebrochen	Ende	September	Morgenstunde
viermonatig	beendet	Badesaison	April	verregnet
fünfmonatig	wiederaufnehmen	Spätherbst	Ostern	Stadthalle
Abschluß	erst	Sendestart	Weihnachtsferien	vormittags
Abschluss	unterbrochen	bezugsbereit	Januar	Kultursaal
achtstündig	unterbrechen	Skisaison	Woche	Familienbildungswerk
dreitägig	geendet	Anfang	März	abgehalten

MK2

Das deutsche Verb *beginnen* wird mit dem Lehnwort *starten* kontrastiert. Es ist offenbar, dass *beginnen* und dessen KP mehr dem KP von *starten* ähneln als dem KP von *anfangen*. Das beweisen die orangenen Farbtöne, die konstant sind und eine breite Überlappung andeuten.

In Quadraten A1 – A3 sind klare KPe von *starten* zu beobachten. Sie umfassen den globalen Kontext einer Raumfahrt (A1 – *Raumfähre*, *Weltraumstation* usw.) oder eines Flugs (A2 – *landen*, *Bord*, *Flug*; A3 – *abfliegen*, *düsen* usw.) Damit hängen auch Quadrate B1 (*Countdown*), B2 (*Testflug*) sowie B3 (*startklar*) zusammen, obwohl sie sich

teilweise mit *beginnen* überlappen. Das Wort *zurückewartet* in B3 entspricht zwar auf den ersten Blick nicht dem Kontext Raumfahrt, trotzdem beweisen einige Belege aus der CCDB, dass das Wort zu diesem Kontext gehören kann. Das KP zeigt nämlich den Kookkurrenzpartner *Cape* und angeführte Beispiele deuten an, dass es sich um „Cape Canaveral“ handelt (ein Gebiet in Florida mit den wichtigsten Weltraumbahnhöfen; www.wikipedia.org, Stand: 1/2014). Den Kontext, den Quadrate A1 – 3 und B1 – 3 darstellen, bestätigen auch die Angaben des Verbs in DUW: „[...] *dass etw. auf ein Ziel hin in Bewegung gesetzt wird*“ (s. Tabelle). Nach Untersuchung der KPe von Wörtern in C1 ist festzustellen, dass sich dieses Quadrat auf den Kontext einer – wahrscheinlich politischen – Kampagne bzw. eines Projektes bezieht. Das Verb *lancieren* bedeutet nach DUW u. a.:

„gezielt in die Öffentlichkeit gelangen lassen (in die Presse); (bes. Werbung, Wirtsch., Politik) einer Sache durch gezielte Maßnahmen zu Anerkennung, Ansehen, Verbreitung verhelfen“

Außer *lancieren* haben ebenso Wörter *initiiieren*, *angelaufen* und *großangelegt* oft den Kookkurrenzpartner *Kampagne* oder *Projekt*, wie weitere Konkordanzen in der CCDB beweisen.

In der linken Ecke unten gibt es vier Quadrate, deren KPe dem KP des Verbs ***beginnen*** am ähnlichsten sind. Die KPe der Wörter im Quadrat D1 (*Beginn*, *einläuten*, *einstimmen*, *angebrochen* usw.) zeigen am häufigsten den Kontext eines Zeitabschnittes wie Ära, Saison, Jahrhundert, Schuljahr, Jahreszeiten, Festtage (sehr häufig beim Verb *einstimmen*: *Ostern*, *Adventszeit* usw.), aber gleichzeitig weisen die KPe ziemlich oft auf politisch-gesellschaftliche Kontexte hin: Wahlkampf, Weltkrieg, Krise u. Ä. Beide Kontexte können deshalb mit dem Verb *beginnen* zusammenhängen. D2 ist das Quadrat, das dem KP von *beginnen* am deutlichsten entspricht. Es beinhaltet Verben, die oft in der Kookkurrenz von Lexemen der Zeitangaben präsent sind. Zwar lässt sich ein konkreter globaler Kontext nicht scharf feststellen, aber wie es in DUW steht, bedeutet *beginnen* einen bestimmten Anfang (Zeit, Ort, Weise; s. Tabelle 1), was dies der Fall sein könnte. Ähnlich sehen Quadrate E1 und E2 aus. E1 beinhaltet antonymische Wörter zum Verb *beginnen* (*beendet*, *unterbrechen* usw.) und E1 einen Kontext der Dauer (*mehrwöchig*, *achtstündig* usw.).

Die **Überlappung** von KPen der zwei Verben ist sehr breit. Sie schließt die folgenden globalen Kontexte ein: Wettkampf, der in DUW bei *starten* erwähnt ist (C2 - *Startschuss, laufen* usw.), Veranstaltung wie ein Kurs (C4 – *Anmeldefrist, Sendetermin, eröffnen* usw., vgl. Tabelle 1) oder wie ein gesellschaftliches Ereignis (E5 – *Stadthalle, Kultursaal* usw.). Bei C4 ist noch zu erwähnen, dass der Kontext ebenso mit Sport zusammenhängen kann (s. u.). Weiter sind hier Quadrate mit Zeitangaben zu beobachten (C5 – Tage, E4 – Monate und E3 – Zeitabschnitte). Genauso wie in der Analyse *anfangen – beginnen*, lässt sich hier der globale Kontext der Arbeit finden (D3; vgl. MK 1 – D2). Die Wörter in Quadrat D4 treten mit Ausdrücken der Zeitangaben und mit ökonomischen Termini vor, wie ihre KPe andeuten. Wichtig ist zu zugeben, dass die Lexeme *Arbeitsbeginn* und *Friedenspflicht* auf einen beruflichen Kontext hindeuten. Die Quadrate D3 – D5 und E3 – E5 weisen auf Ähnlichkeiten zu KPen beider Verben hin, jedoch nicht gleichmäßig, denn der Farbton ist dunkler als orange Farbe, d. h. die KPe stehen näher dem Verb *beginnen*.

Zusammenfassung:

Zusammenfassend lassen sich die folgenden Unterschiede zwischen den Verben *beginnen* und *starten* feststellen: Der globale Kontext bei *starten* ist markant an Flugreise bzw. Weltraumfahrt orientiert. Es geht um eine bedeutsame Reise, die zu bestimmter Zeit geplant wird und die „auf einen Startschuss wartet“ (im Unterschied z. B. zur Reise mit dem Auto u. Ä.). Ein festgelegter Anfang ist stark mit *beginnen* verbunden, allerdings in anderen Zusammenhängen (anhand der MK 2 heißt es vereinfacht „in allen Kontexten außer der Reise“). Dieses Verbpaar ist hinsichtlich der umfangreichen Überlappung in verschiedenen Bereichen sehr synonymisch (v. a. was Zeitangaben betrifft).

4.1.3. Analyse *anfangen* – *starten*

© Cyril Belica: Modelling Semantic Proximity – Contrasting Near-Synonyms (version: 0.21, init tau: 0.4, dist: x, iter: 10000)

starten	anfangen			
Auftaktveranstaltung	Auftakt	gestartet	abfliegen	Start
zurück erwartet	startberechtigt	neuntätig	landen	Weltraumbahnhof
Freitag	ausgetragen	achtetägig	wohlbehalten	MESZ
gastiert	Hauptprobe	halbjährig	Chartermaschine	MEZ
Donnerstag	auslosen	Startplatz	Testflug	planmäßig
Mittwoch	austragen	zehntätig	Landung	abheben
Dienstag	Zahltag	Startrecht	fliegen	Cape
gastieren	drittenmal	Vorbereitungsphase	überfliegen	Raumfähre
Anmeldefrist	Startschuß	angelaufen	startklar	
Bewerbungsfrist	Startschuss	großangelegt	Jungfernflug	
weiterverhandeln	beginnen	lancieren	durchstarten	
weiterverhandelt	eröffnet	lanciert		
verschieben	eröffnen	initiierten		
soweit	fortgesetzt	initiiert		
Anmeldeschluss	fortsetzen	beendet		
Anmeldeschluß	nächste	österreichweit		
anlaufen	kommen	kapiert	angefangen	lernen
begonnen		merken	aufhören	Theaterspiel
planen		ankommen	angewöhnen	verlernt
bereits		kapieren	abgewöhnen	verlernen
ankündigen		losgehen	gewöhnen	Spaß
plangemäß		mitbekommen	Fangen	Tanzen
angekündigt		begriffen	begänne	töpfern
frühestens		wieder	angewöhnt	Fußballspielen
	ja	rauskommen	fangen	fernsehen
	tun	dran	aufwachen	zusammensitzen
	irgendwas	hinschmeißen	los	aufstehen
	schiefgehen	mitkriegen	heimkommen	meditieren
	eingefallen	kriegen	lose	schlafen
	irgend	da	hochkommen	frühstücken
	irgendwann	eh	aufgewacht	Suff
	Aber	reingehen	mittendrin	ausdauern
worüber	gar	rausgehen	hinsetzen	meckern
was	gerne	sowieso	quatschen	kotzen
woran	geschweige	hierherkommen	ficken	lachen
wen	nichts	nachher	reden	atmen
wovon	buchstabieren	jemand	immerzu	schreien
worum	selber	zumute	schwätzen	plärren
wem	niemand	hingehen	dasitzen	heulen
wer	können	irgendwer	scheißen	saufen

MK3

Die letzte Merkmalskarte dieser synonymischen Reihe kontrastiert das deutsche Verb *anfangen* mit dem Lehnwort *starten*. Die MK ist diagonal getrennt und fast ohne Überlappungen. Das bedeutet, dass globale Kontexte dieser Verben sehr unterschiedlich sind.

Die ersten zwei Zeilen entfallen auf das Wort *starten*. Quadrate in der linken Ecke oben (A1 – A2, B1 – B2) weisen auf den Kontext eines sportlichen Ereignisses hin. Diese These wird mit einigen Belegen aus der CCDB und mit Angaben aus dem DUW begründet. DUW führt an, dass *starten* im Kontext des Wettkampfs vorkommt

(s. Tabelle 1). Wettkämpfe finden an bestimmten Tagen statt (A1) und Sportler müssen sich dazu anmelden (B1).

Als Beispiel führe ich folgende Belege aus der CCDB an:

[...] In einer Woche endet die Anmeldefrist für die Champions League [...]

[...] Je näher der Anmeldeschluß für den Europacup kommt, [...]

Den Kontext eines sportlichen Wettkampfes bilden noch Wörter in A2 und ihre KPe bringen Beweise dafür. Nach DUW wird *austragen* u. a. im sportlichen Kontext verwendet (in der CCDB gibt es viele Beispiele mit Kookkurrenzpartnern wie *Rennen*, *Turnier*, *Olympische Spiele* usw.). Vielleicht überraschend kommt das Wort *Zahltag* oft zusammen mit sportlichen Terminen vor:

[...] Großer Zahltag auch für die USA-Staffel [...]

reichen 1:59,93 Minuten vom größten Zahltag der Leichtathletik-Geschichte

Die Pfalz fiebert ihren Feier- und Zahltagen in der Champions League [...]

Es geht höchstwahrscheinlich um einen bedeutsamen Tag, an dem es zu einer positiven Wende kommt. DUW gibt aber keine Informationen zu dieser Bedeutung an.

Quadrat A3 ist eher der rechten Ecke zuzurechnen, die den globalen Kontext des *Flugs* und der *Raumfahrt* zeigt (vgl. MK 2). B3 repräsentiert einen politischen Kontext (*lancieren*, *initiiieren*, *großangelegt* u. a., vgl. MK 2.) und das letzte Quadrat C1 deutet den Anfang einer Veranstaltung oder eines Vorhabens an. Dieses Quadrat ähnelt dem Quadrat C3 aus MK 2, in der die KPe beider Verben (*starten* und *beginnen*) gleichmäßig vertreten sind. Daraus ergibt sich der Kontext eines bestimmten und geplanten Vorhabens, einer Unternehmung oder Veranstaltung, was die typische Verwendung gerade von Verben *beginnen* und *starten* darstellt, allerdings nicht vom Verb *anfangen*.

Die unteren drei Reihen von Quadraten gehören hauptsächlich dem Verb ***anfangen***. Wie schon in der vorherigen Beschreibung (MK 1) angedeutet wurde, sind die globalen Kontexte mit menschlichen Tätigkeiten verknüpft. Das lässt sich bei Quadraten C3 – C5, D4 – D5 und E4 beobachten. Im E5 handelt es sich um menschliche Geräusche, bzw. Ausdrücke (*atmen*, *lachen*, *heulen* u. a., vgl. MK 1). In der linken Ecke befindet sich das

Quadrat mit Fragewörtern (E1 – *was, woran, wovon, wer* u. a.). Dieses syntaktische Merkmal ist in der MK 1 beschrieben, in der sich dasselbe Quadrat (E3) befindet (s. o.).

Die einzigen **Überlappungen** werden mit Quadraten B4 und C2 veranschaulicht. B4 vertritt den Kontext des Flugs. Im C2 steht nur ein einziges Wort, und zwar *kommen*. Trotz der Untersuchung im DeReKo und in der CCDB ist mir keine eindeutige Interpretation gelungen. Eine mögliche Erklärung des Quadrats ist, dass *kommen* für den Kontext der angehenden Zeit bzw. des Zeitabschnittes steht, wie es z. B. in folgenden Sätzen verwendet ist:

Es kommt die Zeit, in der [...]

Der Sommer kommt.

Zusammenfassung:

Die Analyse dieser Merkmalskarte geht von Erkenntnissen aus MK 1 und MK 2 aus. *Starten* und *anfangen* sind unterschiedliche Lexeme, die sich fast nie substituieren lassen. Zum Verb *starten* gehören globale Kontexte wie *Weltraumfahrt, Flug, Sport* und *Politik*, das Verb *anfangen* hingegen ist mit Kontexten wie *menschliche Tätigkeiten, Handlungen* und *Ausdrücke* repräsentiert.

Fazit:

Die Synonymreihe *anfangen* – *beginnen* – *starten*

Es ist bemerkenswert, dass die scheinbar reinen Synonyme *anfangen* und *beginnen* hinsichtlich ihrer KPe markant unterschiedlich sind. Für wichtig halte ich die Erkenntnis, dass *beginnen* in den meisten Kontexten sehr synonymisch zum Lehnwort *starten* vorliegt. Das kann man über *anfangen* nicht sagen, von beiden Verben ist *anfangen* mit seinem KP ziemlich entfernt. Aus stilistischer Sicht sind diejenigen Quadrate sehr auffällig, die den wesentlichen Teil der MK von *anfangen* bilden und die umgangssprachliche, sogar vulgäre Ausdrücke enthalten.

Wenn es eine CNS-Merkmalskarte gäbe, die in der Lage wäre, drei Wörter gleichzeitig miteinander zu kontrastieren, würde es vielleicht nur ein einziges Überlappungsquadrat geben, in dem alle drei Verben vertreten wären. Das würde das Quadrat mit Verb *kommen* betreffen, dessen KP allen drei Verben ähnlich ist. Daraus

folgt, dass der gemeinsame globale Kontext höchstwahrscheinlich ein bestimmter Zeitabschnitt ist.

Die folgenden Tabellen veranschaulichen globale Kontexte der verglichenen Verben und ihre Überlappungen:

BEGINNEN	ANFANGEN
VERANSTALTUNG WETTBEWERB VERLAUF ZEITANGABEN BAUARBEIT VORHABEN	MENSCHLICHES VERHANDELN GESELLSCHAFTLICHE EREIGNISSE KOMMENDER TAG, MORGEN UMGANGSPRACHLICHE VERWENDUNG
TAG/NACHT	

BEGINNEN	STARTEN
ANFANG EINES: <ul style="list-style-type: none"> • ZEITABSCHNITTS • POLITISCHEN/GESELLSCHAFTLICHEN EREIGNISSES NÄHER BESTIMMTER ANFANG von etw.	FLUG WELTRAUMFAHRT KAMPAGNE/PROJEKT
ZEITANGABEN + ZEITABSCHNITTE ARBEIT/BERUF VERANSTALTUNG SPORT	

ANFANGEN	STARTEN
MENSCHLICHE AUSDRÜCKE/AKTIVITÄTEN/TÄTIGKEITEN	FLUG WELTRAUMFAHRT KAMPAGNE/PROJEKT SPORT (WETTKAMPF)
ZEITABSCHNITT	

4.2. Analyse der Synonymreihe: **HERSTELLEN** – **ERZEUGEN** – **PRODUZIEREN**

HERSTELLEN <sw. V.; hat>	ERZEUGEN <sw. V.; hat> [mhd. erziugen, zu ² zeugen]	PRODUZIEREN <sw. V.; hat> [lat. producere (2. Part.: productum) = hervorbringen; vorführen, zu: ducere = führen]
<p>1. <i>gewerbsmäßig in laufender Produktion anfertigen:</i> etw. maschinell, von Hand, synthetisch, billig h.; Autos serienmäßig h.; etw. in Heimarbeit h.; im Ausland hergestellte Produkte.</p> <p>2. a) <i>durch bestimmte Anstrengungen zustande bringen, schaffen:</i> [telefonisch] eine Verbindung, einen Kontakt h.; eine Verbindung zwischen der Insel und dem Festland h.; die Schule möchte ein gutes Verhältnis zum Elternhaus h.; endlich waren Ruhe und Ordnung hergestellt;</p> <p>b) <h. + sich> <i>hergestellt (2. a), erreicht werden:</i> ein so gutes Verhältnis wie früher stellte sich nicht mehr her.</p> <p>3. <i>in den ursprünglichen guten Zustand zurückversetzen:</i> die Kranke war so weit hergestellt (<i>genesen</i>), dass sie aufstehen konnte.</p> <p>4. <i>an den Ort, in die Nähe des Sprechenden stellen.</i></p>	<p>1.</p> <p>a) <i>entstehen lassen; bewirken:</i> Reibung erzeugt Wärme; er versteht es, Spannung zu e.;</p> <p>b) (veraltend) <i>zeugen:</i> er hatte viele Kinder erzeugt.</p> <p>2. a) <i>produzieren, hervorbringen:</i> Strom e.; der Boden erzeugt alles, was wir brauchen;</p> <p>b) (österr.) (<i>Gebrauchsgüter</i>) <i>herstellen:</i> Kleider e.</p>	<p>1. a) (bes. Wirtsch.) <i>erzeugen, herstellen:</i> die Industrie produziert mehr, als sie absetzen kann; Ü (salopp:) Kinder p.;</p> <p>b) <i>für die Herstellung, Finanzierung von etwas sorgen, jmds. Musik-, Filmproduzent sein:</i> eine Compact Disc p.; wer produziert eigentlich diesen Sänger?</p> <p>2. (ugs.) <i>machen; hervorbringen:</i> großen Lärm p.</p> <p>3. <p. + sich> (ugs.) <i>sich [in einer bestimmten Weise] auffallend benehmen [um sein Können zu zeigen]:</i> sich gern [vor anderen] p.; sich als Clown p.</p> <p>4. (bes. schweiz., sonst veraltet) <i>[herausnehmen u.] vorzeigen, vorlegen, präsentieren.</i></p>

Tabelle 2: Angaben in DUW

In der nächsten Analyse werden Beziehungen zwischen Lexemen der synonymischen Reihe *herstellen* – *erzeugen* – *produzieren* erforscht. Es werden (wie in der ersten Analyse) Kookkurrenzprofile und Unterschiede bzw. Übereinstimmungen zwischen zwei deutschen Verben und einem Lehnwort untersucht. Auch bei diesen Lexemen ist es für Nichtmuttersprachler schwierig zu entscheiden, wie man sie richtig verwenden soll, damit es dem realen Sprachgebrauch entspricht.

4.2.1. Analyse *herstellen* – *erzeugen*

© Cyril Belica: Modelling Semantic Proximity - Contrasting Near-Synonyms (version: 0.21, init tau: 0.4, dist: x, iter: 10000)

herstellen	erzeugen			
hervorrufen	generieren	erzeugt	angetrieben	Strom
breitmachen	Entladung	Wärme	Kilowatt	Elektrizität
auslösen	generiert	Wasserstoff	Wirkungsgrad	Stromproduktion
schüren	atmosphärisch	umwandelt	Verbrennungsmotor	Photovoltaik
geschürt	wabern	elektrisch	antreiben	eingespeist
spüren	verflüchtigen	Wärmeenergie	Brennstoffzelle	Blockheizkraftwerk
hervorgerufen	vibrieren	umwandeln	Turbine	Strombedarf
verspüren	transformieren	Kühlwasser	Elektromotor	Kilowattstunde
	ionisieren	verdampfen	Brennstoff	Stromerzeugung
	absorbieren	aufgeheizt	Abwärme	Wasserkraft
	energiereich	emittieren	Erdgas	Biomasse
	absorbiert	entweichen	verbraucht	regenerativ
	Elektron	emittiert	verbrauchen	Windkraft
	Schwingung	Sauerstoff	fossil	Energiegewinnung
	Atom	gasförmig	Kohle	Biogas
	Schallwelle	verdampft	betrieben	Energiequelle
wiederherstellen	gezeugt	freisetzen	produziert	umweltschonend
Wiederherstellung		freigesetzt	produzieren	umweltfreundlich
gewahren		Abfallprodukt	Rohstoff	Gewinnung
gewahrt		zersetzt	weiterverarbeitet	Erzeugung
schaffen		überschüssig	weiterverarbeiten	
Streben		anreichern	liefern	
aufrechterhalten		tiefgekühlt	importieren	
zurückgewonnen			importiert	
herstellbar	Kampfstoff	Grundstoff	Lebensmittel	hochwertig
Verträglichkeit	Giftgas	Vorprodukt	Fertigprodukt	Massenproduktion
getestet	Anthrax	gelagert	Nahrungsmittel	Herstellen
testen	unschädlich	verwenden	Soja	Textilien
entwickeln	Nervengas	lagern	Genußmittel	Textilie
entwickelt	spaltbar	hergestellt	Genussmittel	verkäuflich
maßschneidern	oxidieren	verwendet	exportieren	usw
funktionsfähig	hochgiftig	Lagerung	Einfuhr	Textil
vervielfältigen	gentechnisch	beschichten	wiederverwertbar	fertigen
vervielfältigt	enthalten	Werkstoff	verwendbar	gefertigt
extrahieren	chemisch	Baustoff	Naturprodukt	fabrizieren
Zellkultur	pharmazeutisch	Zellulose	Rohmaterial	beschaffen
injizieren	synthetisch	recycelt	veredelt	fabriziert
patentiert	Zusatzstoff	Dämmstoff	angeliefert	zusammenbauen
züchten	biotechnisch	abbaubar	veredeln	geliefert
injiziert	Ausgangsstoff	Legierung	Weiterverarbeitung	nachbauen

MK4

Das Verb *erzeugen* ist in der größeren Anzahl von Quadraten in der CNS-Merkmalskarte vertreten als *herstellen*. Das lässt sich aus den Farbtönen schließen: Die obere Hälfte der

MK, in der ein konstanter intensiver roter Farbton zu beobachten ist, gehört deutlich dem Verb *erzeugen*. Dagegen zeigt sich die gelbe Farbe der unteren Hälfte nicht so intensiv und ist mit dunkleren Tönen vermischt. Daraus folgt, dass die Felder nicht nur dem Verb *herstellen* gehören und dass man eine markante Übereinstimmung zwischen den Synonymen beobachten kann.

Die KPe der Wörter in den ersten zwei Reihen (A, B) ähneln den KPen des Verbs *erzeugen*. Am stärksten ist ohne Zweifel der Themenbereich Energie- bzw. Stromproduktion vertreten (ein Beispiel ist in DUW unter Lesart 2 angeführt, s. Tabelle 2). Das ist vor allem von Quadraten in der rechten Ecke oben ablesbar. Zwischen den Quadraten bestehen jedoch markante Unterschiede. A5 gibt eine relativ allgemeine Auskunft über den Strombedarf der Zivilisation an. B5 umfasst Lexeme, die auf die umweltfreundliche Stromproduktion hindeuten (*Biomasse, Wasserkraft, Windkraft* usw.) und Quadrat B4 evoziert eher unökologische Energiebenützung (*Kohle, Erdgas* u. a.). Der Kontext der Energie breitet sich bis zur linken Seite der MK aus (B2 – *Atomenergie*). A4 weicht mit seinem globalen Kontext Antrieb vom Thema Energieproduktion ab, allerdings hängt er mit der Energie zusammen. Ganz unterschiedlich ist Quadrat A1 zu verstehen. Aus den beinhalteten Verben geht hervor, dass *erzeugen* im übertragenen Sinne verwendet werden kann. Es lassen sich die folgenden Kookkurrenzpartner dieses Verbs in der CCDB finden: *Empörung, Angst, Panik, Hass, Konflikt, Wut, Schock, Emotionen* u. Ä. Das Wortfeld wird also von Lexemen der subjektiven Gefühle und der mit menschlicher Stimmung verbundenen Abstrakta gebildet. Dies beweisen einige Belege aus der CCDB:

[...]Der enorme Stellenabbau schürt natürlich die Angst.[...]

Das Urteil, das weltweit Empörung auslöste, wurde aufgehoben.[...]

[...]Thema, das gemischte Gefühle hervorruft.[...]

[...]Die Menschen in Bosnien verspüren Sehnsucht nach Demokratie.[...]

Ähnliche Verwendungsweise mit *erzeugen*:

[...]Es sei wichtig, keine Angst zu erzeugen.[...]

Der Kommentar erzeugt das Gefühl, daß [...]

DUW vernachlässigt diese bedeutsame Lesart. Die einzige Angabe, welche die Verwendung im Sinne der Abstrakta bestätigt, lautet „*er versteht es, Spannung zu e.*“ (s. Tabelle 2). In der CCDB hingegen sind zahlreiche Belege mit der abstrakten Anwendung von *erzeugen* zu finden. Aus diesem Grund ist das einzige Beispiel in DUW für Benutzer relativ ungenügend.

Außerhalb des Energiekontextes vertritt noch Quadrat C2 das KP von *erzeugen*, und zwar mit dem einzigen Wort *gezeugt*. Bezüglich der Angabe in DUW (s. Tabelle 2) geht es um eine veraltende Bedeutung, wobei das Wörterbuch folgendes Beispiel anführt: „*er hatte viele Kinder erzeugt*“. Das Wort *erzeugt* stimmt zwar nicht mit der Form in Quadrat C2 überein, aber die CCDB beinhaltet mehrere Belege, in denen gerade *gezeugt* präsent ist, z. B.: [...]König Georg III. soll zwei Söhne gezeugt haben[...]

Aus dem Rest der MK ist ersichtlich, dass sich die Verben zum großen Teil gegenseitig **überlappen**. Reihe C (außer Quadrat C1) neigt mehr dem Verb *erzeugen* zu. Es handelt sich um Kontexte der Gewinnung von Rohstoffen und damit verbundenes Geschäft und Produktion (C4 – *produzieren, liefern, importieren* u. a.; C5 – *Gewinnung* u. a.). Die unteren Zeilen beziehen sich auf verschiedene Branchen im Bereich der Industrie. Die linke Ecke unten stellt Kontexte aus dem Chemie-Bereich dar. E1 und E2 evozieren Experimente im Labor und Herstellung von Medikamenten, d. h. die pharmazeutische Industrie. Es stellt sich die Frage, ob es bei Quadrat D2 um Waffenindustrie geht. Offensichtlich sind hier aber chemische Waffen thematisiert. Quadrat E3 hängt mit der Papier- bzw. Textilindustrie zusammen, wie sich aus Wörtern wie *Zellulose* oder *recycelt* schlussfolgern lässt. Ohne Zweifel ist der globale Kontext Textilindustrie in D5 dargestellt. Das benachbarte Quadrat D4 weist auf Lebensmittelindustrie hin. Insgesamt sind diese Kontexte an Gebrauchsindustrie orientiert und bezüglich ihrer KPe nähern sich die unteren Quadrate eher zum Verb *herstellen* an. Das entspricht der Angabe in DUW: „*gewerbsmäßig in laufender Produktion anfertigen*“ (s. Tabelle 2).

Das einzige Quadrat, welches das Verb ***herstellen*** ohne Überlappungen repräsentiert, ist C1. Im Unterschied zu anderen Kontexten handelt es sich hier um keine materielle Herstellung. DUW führt mehr Beispiele unter der Unterbedeutung Nr. 2 an, die dem Quadrat entsprechen (s. Tabelle 2). Die Erklärung lautet: „*durch bestimmte*

Anstrengungen zustande bringen, schaffen“. Der globale Kontext in C1 könnte wie folgt verstanden werden: Bemühung, höhere menschliche Werte zu pflegen. Unter häufigen Konkurrentenpartnern findet man in der CCDB Wörter wie: *Frieden, Kontakt, Beziehung, Freiheit* usw. Genauso wie es bei *erzeugen* der Fall war (A1), geht es hier um Abstrakta. Ich führe wieder Belege aus der CCDB als Beispiel an, in denen Wörter aus C1 und das Verb *herstellen* auftreten:

So könnten Einheit und Kontinuität gewahrt bleiben.[...]

[...]Es symbolisiert das Streben nach Unabhängigkeit und Freiheit[...]

[...]Gerechtigkeit schafft Frieden[...]

[...]Die persönliche Beziehung ist hergestellt, die Atmosphäre locker. [...]

[...]Kontakte zu anderen Menschen herstellen[...]

Zusammenfassung

Beide Wörter verfügen über einen bestimmten Themenbereich, der für sie typisch ist. Für *erzeugen* ist es das Thema Energie, für *herstellen* Gebrauchsindustrie. Man darf allerdings nicht vergessen, dass *erzeugen* in den meisten Kontexten eine große Konkurrenz für *herstellen* darstellt, denn *erzeugen* kann fast in allen Kontexten *herstellen* ersetzen, umgekehrt ist das jedoch nicht möglich. Beide Verben treten u. a. in Kontexten außerhalb des Themenbereichs der materiellen Produktion auf. In diesen Fällen haben sie jeweils einen eigenen Kontext und die Verben können einander nicht vertreten. Mit *erzeugen* sind meistens solche Lexeme verbunden, die negative Konnotationen (Merkmale) signalisieren (obwohl man in der CCDB auch solche Belege finden kann, die positive Merkmale ausdrücken – jedoch nicht in so starker Vertretung). Die abstrakte Verwendung von *herstellen* wirkt eher neutral und kommt am häufigsten in Zusammenhang mit gesellschaftlichen Beziehungen vor.

4.2.2. Analyse herstellen – produzieren

© Cyril Belica: Modelling Semantic Proximity – Contrasting Near-Synonyms (version: 0.21, init tau: 0.4, dist: x, iter: 10000)

produzieren	herstellen			
Fernsehproduktion	vermarktet	zugekauft	eingespeist	wiederherstellen
synchronisieren	vermarkten	zukaufen	Ökostrom	Wiederherstellung
gesendet	wirtschaften	einkaufen	Windkraftanlage	gewahren
Werbespot	exklusiv	beliefern	Atomstrom	gewährt
synchronisiert	Produzieren	verkaufen	Solarstrom	schaffen
herausbringen	arbeiten	Qualitätsprodukt	Biogas	Streben
gedreht	meistverkauft	anbieten	Biomasse	aufrechterhalten
ausstrahlen	kommerziell	Markenartikel	Jahresproduktion	zurückgewonnen
	Massenproduktion	exportiert	verbrauchen	erzeugen
	importiert	abfüllen	verbraucht	erzeugt
	importieren	angebaut	umweltschädlich	gasförmig
	geliefert	Speiseeis	Erdöl	Abwasser
	verkäuflich	abgefüllt	Brennstoff	ungiftig
	beschaffen	Milchprodukt	Tonne	organisch
	Produktion	Transportweg	Kunstdünger	Kohlendioxid
	exportieren	anbauen	Öl	Wasserstoff
gefertigt	usw	weiterverarbeiten	produziert	hemmen
fertigen	Textil	verarbeiten	angereichert	Botenstoff
fabrizieren	Kosmetikum	Rohstoff	Abfallprodukt	lebensnotwendig
fabriziert	Genußmittel	verarbeitet	überschüssig	schütten
zusammenbauen	Genussmittel	weiterverarbeitet		Kohlenhydrat
beschafft	Kinderspielzeug	Naturprodukt		Stoffwechsel
verschrotten	Textilien	gelagert		Rezeptor
angefertigt	Textilie	Fertigprodukt		Nährstoff
verwendbar	Zusatzstoff	hergestellt	pflanzlich	synthetisieren
Baustoff	Soja	Vorprodukt	enthalten	freisetzen
Verarbeitung	Anbau	Herstellung	tierisch	freigesetzt
textil	naturbelassen	Grundstoff	gentechnisch	injizieren
handelsüblich	Zuckerrübe	Endprodukt	biotechnologisch	injiziert
Verpackungsmaterial	Waschmittel	Lagerung	chemisch	körpereigen
Schmierstoff	Baumwolle	pharmazeutisch	biologisch	synthetisiert
Garn	Konservierungsstoff	Ausgangsstoff	löslich	Interferon
beschichten	verwenden	Giftgas	Herstellen	extrahieren
Werkstoff	synthetisch	Nervengas	biotechnisch	vervielfältigen
recycelt	verwendet	Anthrax	neuartig	Zellkultur
Dämmstoff	beigemischt	hochgiftig	Verträglichkeit	genetisch
Legierung	spaltbar	Gefechtskopf	getestet	Labor
kompostierbar	Zellulose	Blausäure	experimentiert	Genmanipulation
wiederverwertet	beimischen	vernichten	patentiert	Erbinformation
keramisch	abbaubar	Kampfstoff	testen	vervielfältigt

MK5

Verben *produzieren* und *herstellen* sind in vielen Kontexten austauschbar, wie die konstanten orangen Farbtöne in der MK andeuten. Die Kookkurrenzprofile dieser Verben ähneln einander, nur in zwei Fällen sind sie ganz unterschiedlich. Die Analyse dieser MK geht von den Erkenntnissen der vorherigen Analyse aus, denn die Quadrate und ihre Wörter wiederholen sich.

Der größte Teil der MK veranschaulicht die **Überlappung**. Es sind einige themenspezifische Felder festzustellen. Einen beträchtlichen Teil bildet der Themenbereich Verbrauchsindustrie (C1 – C4, D1 – D3, E1 – E2), vgl. MK 4. Es sind globale Kontexte

verschiedener Industriezweige zu betrachten – z. B. Textilindustrie (C2, D1), Lebensmittelindustrie (D2) oder Papierindustrie (E1 – E2). Quadrate in der rechten Ecke unten deuten ein naturwissenschaftlich gezieltes Gebiet – wahrscheinlich den Chemie-Bereich und die damit verbundene Pharmazeutik (vgl. MK 4). Quadrat E3 repräsentiert chemische Waffen (vgl. MK 4), C5 biologische Vorgänge in Organismen (*Stoffwechsel, Nährstoff* u. a.). Je höher die Quadrate liegen, desto heller ist der Farbton, d. h. die KPe neigen mehr zu *produzieren*. A2 – A3 sowie B2 – B3 signalisieren den globalen Kontext des Handels mit Produkten (Beispiele: A2 – *meistverkauft*; A3 – *einkaufen, beliefern*; B2 – *exportieren*; (Massen)Produktion und Lebensmittelproduktion in B3 – *Milchprodukt*). Eine neue Information bringt Quadrat B4. Wörter wie *Erdöl, Brennstoff, umweltschädlich* oder der aus Erdöl hergestellte *Kunstdünger* bilden den globalen Kontext der petrochemischen Industrie. Dieser Zweig der Industrie bleibt in der MK 4 aus.

Eine spezifische Bedeutung von ***produzieren*** ist in Quadrat A1 zu beobachten. Sein globaler Kontext deutet klar an *Finanzierung von etwas* hin, was meistens mit Medien oder medialer künstlerischer Richtung zusammenhängt (A1 – *Fernsehproduktion, gedreht, gesendet, Werbespot* u. a.). Die Erklärung dieser Bedeutung ist auch anhand der Angabe in DUW untermauert (s. Tabelle 2). Um einen Farbton dunkler ist Quadrat A4, doch kann es noch zu *produzieren* zugeordnet werden. Der Kontext ist anscheinend *Stromproduktion*, genauso wie es bei *erzeugen* in MK 4 (B5) veranschaulicht ist. Daraus könnte man schließen, dass *produzieren* und *erzeugen* in dem Zusammenhang mit Stromproduktion frei austauschbar sind. Wenn man sich aber die folgende Merkmalskarte (MK 6, z. B. E4 – E5) ansieht, kann man daran zweifeln, denn dieser Kontext kommt trotz einiger Überlappungen offensichtlich mehr zu *erzeugen* nahe (s. u.).

Das einzige rote Quadrat, also der Ausdruck des Verbs ***herstellen***, vertritt die Bedeutung von „Beziehungen anknüpfen und pflegen“ (Erklärung s. MK 4).

Zusammenfassung:

Die Verben *produzieren* und *herstellen* sind mit Bezugnahme auf die Merkmalskarte sehr nahe Synonyme. Die größte Diskrepanz beruht auf der Bedeutung von medialer Produktion und ihrer Realisation, die zu *produzieren* gehört. *Herstellen* unterscheidet sich in der abstrakten Verwendung. In anderen Kontexten sind aber die beiden Verben austauschbar.

4.2.3. Analyse erzeugen – produzieren

© Cyril Belica: Modelling Semantic Proximity – Contrasting Near-Synonyms (version: 0.21, init tau: 0.4, dist: x, iter: 10000)

produzieren	erzeugen			
hemmen	freisetzen	ionisieren	generieren	hervorrufen
Botenstoff	freigesetzt	energiereich	Entladung	breitmachen
vervielfältigen	synthetisieren	Elektron	generiert	auslösen
injiziert	Photosynthese	Schwingung	atmosphärisch	schüren
synthetisiert	Molekül	Atom	vibrieren	geschürt
Eiweißstoff	zersetzt	Schallwelle	transformieren	spüren
injizieren	anreichern	elektromagnetisch		hervorgerufen
Stoffwechsel		hochfrequent		verspüren
enthalten	hergestellt	absorbieren	umwandelt	speien
lebenswichtig	herstellen	absorbiert	verdampfen	wabern
Eiweiß	Grundstoff	emittieren	aufgeheizt	verflüchtigen
lebensnotwendig	umweltschädlich	unempfindlich	entweichen	verpuffen
schütten	Lagerung	Plasma	emittiert	verpufft
Kohlenhydrat	hochgiftig	gezeugt	Sauerstoff	
tierisch	pharmazeutisch	bestrahlt	verdampft	
Anthrax	biotechnologisch	filtern	umwandeln	
Vorprodukt	exportieren	überschüssig	Wärme	angetrieben
Naturprodukt	exportiert	Rohstoff	heizen	elektrisch
Erzeugnis	geliefert	Abfallprodukt	Gas	Verbrennungsmotor
Fertigprodukt	gelagert	gasförmig	Kühlwasser	antreiben
Agrarprodukt	tonnenweise	Treibstoff	Sonnenstrahlung	Turbine
Getreide	lagern	liefern	versorgt	Elektromotor
Nahrungsmittel	abfüllen	Kohlendioxyd	umgewandelt	Generator
Genußmittel	beschaffen	Methanol	versorgen	Gasturbine
eingekauft	produziert	erzeugt	Elektrizität	Blockheizkraftwerk
zugekauft	Jahresproduktion	Erzeugung	fossil	Abwärme
Produkt	Jahreskapazität	umweltfreundlich	Kohle	Megawatt
Massenproduktion	Produktionskapazität	verbraucht	Primärenergie	Kilowatt
zukaufen		verbrauchen	beheizt	Wärmepumpe
Produktion		Atomstrom	Gewinnung	Warmwasser
Massenprodukt		Brennstoff	beheizen	speisen
einkaufen		Biogas	Heizen	Wirkungsgrad
gefertigt	Fernsehproduktion		Stromerzeugung	Windkraft
fertigen	synchronisieren		regenerativ	Strombedarf
fabrizieren	gesendet		Energiequelle	Einspeisung
fabriziert	Werbespot		Energieträger	eingespeist
verkauft	synchronisiert		erneuerbar	Wasserkraftwerk
vermarktet	herausbringen		Energieerzeugung	kWh
vermarkten	gedreht		Stromgewinnung	Ökostrom
wirtschaften	ausstrahlen		Energieform	Stromversorgung

MK6

Die letzte Kontrastierung der Dreiergruppe synonymischer Verben stellt das deutsche Verb *erzeugen* und das Lehnwort *produzieren* gegeneinander. Genauso wie in vorigen Merkmalskarten kann man einen hohen Ähnlichkeitsgrad von KPen dieser Verben beobachten, denn der Farbton ist in der ganzen Merkmalskarte sehr stabil.

Die linke Hälfte der MK neigt eher zu *produzieren*. Jedoch gehört nur ein Quadrat ausschließlich diesem Verb, und zwar E2, dessen globaler Kontext Finanzierung von der medialen Produktion aufzeichnet (die Erklärung ist in MK 5 näher beschrieben).

Die anderen Quadrate der linken Seite stellen eine beachtliche **Übereinstimmung** dar. Die sich in der unteren Ecke befindenden Quadrate vertreten den wirtschaftlichen Kontext. Das Thema ist an Industrie, Produktion von Waren und Handel gerichtet (vgl. MK 4 und MK 5). Der obere Bereich rechts umfasst den Kontext der naturwissenschaftlichen Gebiete wie *Biologie* und *Chemie*. A1 – A2 und B1 illustrieren mit ihren Wörtern biologische Prozesse und für lebendige Organismen notwendige Stoffe (A1 – *Eiweißstoff*, *Botenstoff* u. a.; B1 – *lebenswichtig*, *Kohlenhydrat* u. a.). Die rechte Seite der MK kommt dem Verb *erzeugen* näher, stellt aber stets gewissermaßen die Überlappung dar. Quadrate A4, B3 und B4 gehören zum *Physik-Bereich*. B5 evoziert Erscheinungen, welche die Luftbewegung begleiten. In der rechten Ecke unten sind Quadrate mit dem Kontext Energie konzentriert. Es sind sowohl die Energieerzeugung (D3 – D4, E4 – E5), als auch ihr Verbrauch (C5, D5) vertreten (mehr dazu in MK 4). Wörter in Quadrat C3 entsprechen dem Kontext der Ausnutzung von Treibstoffen (*Treibstoff* oder *Methanol*, das als Kraftstoff eingesetzt wird) und ihre Konsequenzen (*Kohlendioxid*, *Abfallprodukt*).

Genauso wie die anderen zwei Verben repräsentiert das Verb ***erzeugen*** nur ein Minimum von Quadraten ohne Überlappungen (A3 und A5). A5 beinhaltet Wörter, die oft in übertragener Bedeutung verwendet werden (s. MK 4). A3 signalisiert das Physik-Fachgebiet.

Zusammenfassung:

Die Diskrepanz zwischen *produzieren* und *erzeugen* ist nur sehr gering. Sie beruht auf Kontexten medialer Produktion (*produzieren*) und der spezifischen Bedeutung des Verbs *erzeugen*, die im Zusammenhang mit menschlichen (oft negativen) Gefühlen auftaucht und derer Sinn abstrakt ist. Die Überlappung bildet ein sehr umfangreiches Feld. Es beinhaltet solche Kontexte, die in den vorherigen Merkmalskarten erwähnt und beschrieben wurden (*Stromerzeugung*, *Industrie*, *Chemie- und Physikbereich*, *Biologie* u. a.).

Fazit:**Die Synonymreihe *herstellen* – *erzeugen* – *produzieren***

Die Bedeutungsnuancen zwischen diesen drei Verben sind viel geringer als bei der ersten synonymischen Gruppe (MK1 – MK3). Zusammenfassend ist festzuhalten, dass es nur wenige Kontexte gibt, in denen diese Verben nicht frei austauschbar sind und die lediglich einem bestimmten Verb gehören. Das heißt, dass sich die Verben breit überlappen. Der größte Unterschied ist beim Wortpaar *herstellen* – *erzeugen* (MK 4) zu beobachten. Mit dem Fremdwort hingegen kann man beide deutschen Verben fast in allen Themenbereichen austauschen. Es existieren im Prinzip lediglich drei globale Kontexte, die ausschließlich nur zu einem Verb zu zählen sind, wie es in den folgenden Tabellen deutlicher veranschaulicht ist. Es handelt sich um Kontexte der gesellschaftlichen Beziehungen (*herstellen*), subjektiven Gefühle (*erzeugen*) und der Medienproduktion (*produzieren*). Diese Analyse bringt noch eine Erkenntnis: In keiner der Merkmalskarten ist die umgangssprachliche Lesart des Verbs *produzieren* festzustellen (im Sinne von einem „auffallenden Benehmen“), wie es DUW anführt (s. Tabelle 2). Entweder ist diese Verwendung im Sprachgebrauch nicht besonders üblich, oder der Grund ist im Korpus zu suchen, das über gesprochene Sprache nur begrenzt verfügt.

HERSTELLEN	ERZEUGEN
GESELLSCHAFTLICHE BEZIEHUNGEN	ENERGIE STROMPRODUKTION NACHKOMMEN SUBJEKTIVE GEFÜHLE
CHEMIE GEBRAUCHSINDUSTRIE	

HERSTELLEN	PRODUZIEREN
BEZIEHUNGEN	MEDIENPRODUKTION STROMERZEUGUNG
INDUSTRIE WIRTSCHAFT BIOWAFFEN NATURWISSENSCHAFT (Experimente im Labor) BIOLOGISCHE PROZESSE ERDÖL + BENUTZUNG	

ERZEUGEN	PRODUZIEREN
übertragener Sinn - SUBJEKTIVE GEFÜHLE PHYSIK	MEDIENPRODUKTION
BIOLOGISCHE PROZESSE WIRTSCHAFT STROMERZEUGUNG ENERGIEVERBRAUCH PHYSIK	

5. Praktische Anwendung der CCDB

Die CCDB ist eine wertvolle Quelle für Übersetzer, Linguisten und insbesondere für Lexikografen, die aus ihren Daten relevante Erkenntnisse über Lexembeziehungen schöpfen können. Neben der (theoretischen) Linguistik können die Merkmalskarten auch im Sprachunterricht zur Geltung kommen. Die nächsten zwei Unterkapitel sind dem praktischen Gebrauch der CCDB gewidmet, und zwar im Bereich der Lexikografie und Didaktik (DaF-Unterricht).

5.1. Lexikografie

Für lexikografische Untersuchungen sind alte Methoden nicht mehr ausreichend, weil bei ihnen das zu erforschende Material sehr beschränkt ist und die Schlussfolgerungen aus diesem Grund unvollständig sind. Die Arbeit mit alten Wörterbüchern und bescheidenen Glossaren genügt nicht. Korpora bedeuten für die Lexikografie eine neue Alternative und gleichzeitig heutzutage die beste Möglichkeit, Lexeme zu untersuchen. Sie ermöglichen es, den authentischen Sprachgebrauch auszuwerten und die entstandenen lexikografischen Werke, die auf der Basis von Korpora erarbeitet werden, können den Benutzern ausführliche Angabe über den Sprachgebrauch anbieten. Außerdem verfügen Korpora über neue Tools und Funktionen, die die jeweiligen Analysen Untersuchung erleichtern können. Auf der anderen Seite ist auch ein Nachteil der Korpora zu nennen: Sie beinhalten zwar authentische Diskurse, aber wie im Kapitel 3 erwähnt ist, handelt es sich meistens um die geschriebene Sprache (sehr oft Zeitungsartikel). Die konzeptionelle Schriftlichkeit und Mündlichkeit sind wohl nicht gleichmäßig vertreten, was man im Hinterkopf haben muss und was sich in der Forschung widerspiegelt.

Synonymie ist in der Lexikografie immer noch unzulänglich dargestellt. Die meisten Synonymwörterbücher stellen lediglich Wortlisten ohne wichtige semantische und syntaktische Beziehungen zur Verfügung. Als Beispiel führe ich Angaben aus dem zweisprachigen Wörterbuch FIN Publishing (2005):

an|fangen* (i, h. a) začít, počínat; **mit ihm ist nichts anzufangen** (on) není k ničemu
beginnen* (a, h. o) začít, začínat, počít; začínat (se) ANT (*be*)**enden, beschließen**

Aus den Angaben ist ersichtlich, dass die tschechischen Sprecher nicht in der Lage sind, anhand von diesen Informationen eine relevante Folgerung zu ziehen.

Die Überlappungen können in Wörterbüchern nicht visualisiert dargestellt werden, wie es bei Merkmalskarten der Fall ist und es fehlen Beschreibungen von paradigmatischen Sinnrelationen zum großen Teil. Die Arbeit mit dem Korpus ist für Lexikografen zwar mühsam und zeitaufwendig, aber sehr bedeutsam. Wie schon erwähnt, ermöglichen Korpora die Relationsstrukturen von Lexemen zu analysieren und kontextuelle Variabilität zu erfassen, was bei der Herstellung eines Wörterbuches von großer Bedeutung ist (vgl. Storjohann 2006: S.1).

Zu nützlichen Tools gehören auch SOM oder CNS, die für lexikografische Anwendung überaus fruchtbar sind. Es werden einerseits Prinzipien von der Eingliederung in Wortfamilien (Clustering) und Verbindungen von Stichwörtern untersucht (mithilfe von SOM), andererseits die Disambiguierung von Synonymen (CNS). Weiterhin werden Kollokationen und der Phraseologiebereich erforscht (vgl. Vachková 2007b: S. 11-15). Lexikografen können anhand von Merkmalskarten vor allem solche Lexembeziehungen überprüfen, die lange vernachlässigt wurden oder verborgen waren. Die Kookkurrenzanalyse bildet die Grundlage für die Strukturierung des Lexems und seine Bedeutung. Da die Kookkurrenzanalyse (bzw. Merkmalskarte) verschiedene Kontexte des Lexems präsentiert, lassen sich mehrere Übersetzungsmöglichkeiten in unterschiedlichen Kollokationen und syntaktischen Positionen feststellen, was natürlich v. a. für die Erstellung bilingualer Wörterbücher von wesentlicher Bedeutung ist. Das Hauptziel ist es, Synonyme bzw. Hyponyme zu Lexemen zu finden, was mithilfe von MKen möglich ist. Die festgestellten Synonyme fördern die lexikografische Arbeit eines Übersetzungswörterbuches (vgl. Vachková 2007b: S. 110ff).

Ein großer Vorteil der Merkmalskarten ist u. a. die Gewährung einer sehr schnellen, ersten Orientierung, was die Interpretation eines Lexems betrifft.

So ist zum Beispiel beim Blick auf die CNS-Merkmalsskarten gleich ersichtlich, ob die kontrastierten Wörter sehr synonymisch sind oder ob sich die Kontexte beider Wörter stark unterscheiden.

Die SOM-Merkmalsskarten hingegen umfassen leere oder mehrfarbige Quadrate, welche andeuten, dass das zu erforschende Lexem mehrere Bedeutungen hat.

Die lexikografische Arbeit ausschließlich mit Korpora genügt natürlich nicht. Bei der Analyse von Korpusbelegen müssen ältere Informationen aus Wörterbüchern berücksichtig-

sichtigt werden. Ein großer Anteil von journalistischen Texten innerhalb der Korpora, die über solche Typen von Texten in hohem Maß verfügen, kann die Bedeutungen, die am häufigsten vorkommen, entstellen. Deshalb sollte die Untersuchung aus beiden Quellen – Wörterbüchern und Korpora – schöpfen (vgl. Vachková 2007a: S. 14).

Weiter ist zu erwähnen, dass digitalisierte lexikografische Werke (z. B. auf der CD-ROM) mit gedruckten Wörterbüchern konkurrieren. Digitalisierte Wörterbücher verfügen über neue Möglichkeiten und Funktionen, die den Tools von Korpora ähneln. Die Arbeit mit solchen Wörterbüchern wird damit erleichtert und beschleunigt. Ein digitalisiertes Wörterbuch kann z. B. Hinweise beinhalten, die das Angebot von Informationen erhöhen und sehr oft geht es von Sprachdaten aus Korpora aus. Solche Wörterbücher verfügen deshalb über authentische Beispiele aus dem realen Sprachgebrauch, was für Benutzer günstig ist (vgl. Vachková 2007b: S. 157).

Zusammenfassend kann man ergänzen, dass es zwar nie möglich sein wird, gesamte Kontexte, in denen die Lexeme erscheinen, in Wörterbücher einzugliedern, dennoch können zahlreiche Unvollständigkeiten mithilfe der CCDB beseitigt werden.

5.2. Anwendung im DaF-Unterricht

Eine Merkmalskarte beinhaltet Gruppen von Wörtern, die als Grundlage für den Unterricht dienen können. Da die Merkmalskarten über authentische, themenbezogene und mannigfaltige Wörter verfügen, sind sie vor allem beim Wortschatzlernen äußerst hilfreich. Die Beziehungen zwischen Wörtern, die der Studierende betrachtet, erleichtern es, die Vokabeln beim Lernen ins Langzeitgedächtnis zu übertragen. Beim Lernen brauchen die Studierenden vielfältige Beispiele, um die Zusammenhänge zwischen den Wörtern zu erkennen. Zugleich entsteht ein Assoziationsnetz zu verschiedenen Kontexten, Inhalten oder Situationen. So werden die neuen Vokabeln im Gedächtnis fest verankert (mehr dazu in Storch 1999: S. 55ff). Die Arbeit mit Merkmalskarten ist hauptsächlich für Fortgeschrittene geeignet (nach meiner Einschätzung etwa ab Niveau B2 des europäischen Referenzrahmens für Sprachen), weil die MKen einen breiten Wortschatz umfassen und Anfänger damit Schwierigkeiten hätten. Die Schüler müssen über einen großen Wortschatz und außersprachliche Kenntnisse verfügen, damit sie die grundsätzlichen Lexembeziehungen entdecken können. Für unerfahrene Studierende wäre eine Arbeit mit MKen zu anstrengend, chaotisch und deshalb demotivierend.

Beim Lernen arbeiten beide Hemisphären des Gehirns und mit ihren unterschiedlichen Funktionen ergänzen sie einander. Die linke Hemisphäre ist für das logische Denken verantwortlich, die rechte hingegen verarbeitet neue Informationen und steht in Verbindung mit Kreativität, Fantasie, räumlich-visueller Orientierung usw. (vgl. Storch 1999: S. 41). Die visuelle Form der SOM-Merkmalskarte ist dabei sehr nützlich, denn sie unterstützt die Zusammenarbeit beider Gehirnhälften. Das mithilfe von Wortvernetzungen aktive Lernen verläuft effektiver als die Aneignung der abgetrennten, isolierten Vokabeln, wie es z. B. in einem Wortregister der Fall ist (vgl. Marková 2012b: S. 62).

Im Rahmen der Unterrichtsarbeit können verschiedene Fertigkeiten anhand der Merkmalskarten geübt werden: Wortschatzlernen, Übungen zur Morphologie und Syntax, mündlicher Ausdruck (Diskussion zu einem Thema) und nicht zuletzt schriftlicher Ausdruck. Die Arbeit mit den Merkmalskarten bietet eine hervorragende Gelegenheit, kreatives Schreiben zu entwickeln, denn sie können als ein Gerüst für Aufsätze dienen (vgl. Vachková et al. 2008: S. 22). Gerade die Produktion eines schriftlichen Textes verhilft den Lernenden dazu, neue Vokabeln, Konstruktionen u. Ä. besser im Gedächtnis zu speichern als die gesprochene Produktion, die eher spontan und ohne Planung verläuft. Das Schreiben ist ein lernintensives Medium, weil die visuellen und schreibmotorischen Eindrücke beim Lernen schneller wirken (vgl. Storch 1999: S. 214).

Außer den angeführten Fertigkeiten ist die Arbeit mit den MKen auch zur Vermittlung von Informationen über Landeskunde geeignet. Es lassen sich Wörter in der CCDB eingeben, die mit Kultur, Geografie u. a. zusammenhängen (z. B. Dresden, Bundesland, Wagner, Weimar, Berlinale usw.), und die Schüler können mit SOM-Merkmalskarten arbeiten und Bedeutungen, Beziehungen und Hinweise zum Kontext suchen. Zur Vermittlung von Landeskunde können natürlich auch die CNS-Merkmalskarten angewendet werden. Es lassen sich z. B. die folgenden Wortpaare kontrastieren: Berlin-München, BRD-DDR, Deutschland-Österreich usw. Die Studierenden können anhand der MKen über Unterschiede und Ähnlichkeiten diskutieren.

Ein Vorteil der MK ist, dass sie mithilfe von kontextuell organisierten Clustern im Grund einem Assoziogramm (sog. Mind-Map) ähnelt. Das Assoziogramm kann als Grundlage für Sprachproduktion dienen und anhand von seinen thematisch spezifischen Bereichen lässt es sich für Verarbeitung ausgewählter Themen nutzen. Gleichzeitig stellt eine MK systematisierte Wortfelder mit Vokabeln zur Verfügung, die für das Assoziogramm passend sind.

Die Unterrichtsvorbereitung kann jedoch Schwierigkeiten mit sich bringen. Die Arbeit mit der MK sollte dem Niveau der Schüler angemessen sein. Der Lehrer muss zunächst allein die MK analysieren und Aufgaben erstellen, was sehr zeitaufwendig ist. Darüber hinaus sind nicht alle MK schnell zu entschlüsseln – das ist nicht nur für Studierende problematisch, sondern auch für den Lehrer. Trotzdem lassen sich die Merkmalskarten (sowohl CNS als auch SOM) im DaF-Unterricht verwenden. Am Beispiel des folgenden selbst entwickelten Arbeitsblatts möchte ich zeigen, wie Aufgaben für den DaF-Unterricht aussehen könnten.

6. Arbeitsblatt zu MK 1 (CNS *anfangen* – *beginnen*): Ein Entwurf

Niveau B1-B2 (höhere Mittelstufe)

Die Wörter *anfangen* und *beginnen* sind nahe Synonyme. Die CNS-Merkmalskarte kontrastiert ihre Kontexte, in denen sie verwendet werden. Schauen Sie sich die Merkmalskarte an und versuchen Sie einige Themenbereiche zu bestimmen.

- **Mündlicher Ausdruck:**

Aufg. 1: Wie hängen Quadrate A5, B5 und C3 zusammen?

➤ Wie sieht Ihr Morgen aus?

Aufg. 2: Sehen Sie sich Quadrat B2 an.

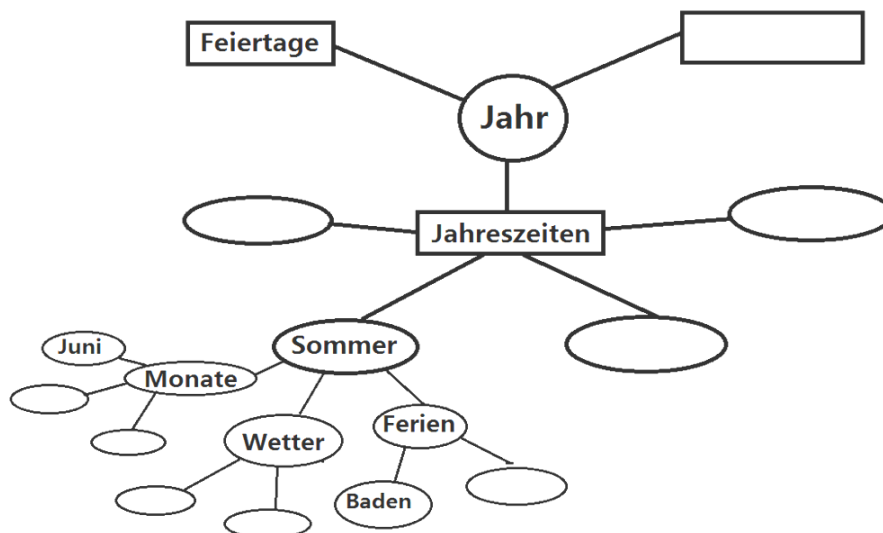
Erweitern Sie es um andere semantisch verwandte Wörter und zeichnen Sie eine Mind-Map. Sie können dabei Wörter aus der Merkmalskarte und eigene Kenntnisse benutzen.

Hilfe!

Bereiche, welche die Mind-Map umfassen kann:

Jahr, Jahreszeiten, Monate, Feste, Saisons, Wetter, Sportarten...

Beispiel:



- Vergleichen Sie Ihr Assoziogramm mit anderen Mitschülern/Komillitonen.
- Besprechen Sie jetzt das Thema *Jahresablauf*. Benutzen Sie dazu das von Ihnen entworfene Assoziogramm.

Aufg. 3: Diskutieren Sie Wörter in Quadrat C5.

Um welches Thema geht es in diesem Quadrat?

Hilfe!

Woran haben Sie Spaß?

Haben Sie diese Tätigkeiten schon probiert?

Was kann man nicht verlernen?

- **Schriftlicher Ausdruck:**

Aufg. 4: Schreiben Sie einen Absatz zum Thema Angewohnheit. Nehmen Sie Stellung zu diesem Thema. Verwenden Sie alle Wörter aus Quadrat C4. Sie dürfen ihre Wortformen verändern.

Hilfe!

Was kann/muss man sich angewöhnen?

Womit soll/möchte man aufhören?

Was ist eine positive und negative Angewohnheit?

Niveau B2-C1 (Fortgeschrittene)

Aufg. 5: Sehen Sie sich Quadrat E1 an. Wie lang kann ein Zeitabschnitt sein?
Schreiben Sie so viele Wörter wie möglich auf:

mehrwöchig, fünfmonatig, unendlich, _____

➤ Wodurch unterscheiden sich Wörter *monatig* und *monatlich*?

Fügen Sie weitere ähnliche Beispiele ein:

Aufg. 6: Schauen Sie sich die Merkmalskarte an.

➤ Welche Quadrate beinhalten umgangssprachliche Ausdrücke?

➤ Welche Lexeme sind das? Gehören sie zum Verb *anfangen* oder *beginnen*?

Aufg. 7: In welchem Quadrat finden Sie onomatopoetische Verben?

- Stellen Sie sich die Geräusche vor und sagen Sie laut Interjektionen zu diesen Verben. Vergleichen Sie Ihre Ideen miteinander.

Beispiel: *zischen*: „tsssss“

- Welche anderen onomatopoetischen Verben kennen Sie?
- Welche sind mit Menschen und welche mit Tieren bzw. Sachen verbunden?
- Vergleichen Sie diese Verben mit Verben Ihrer Muttersprache. Sind Sie ähnlich?

Aufg. 8: Mit welchem Thema könnten Quadrate C2 und D2 zusammenhängen?
Diskutieren Sie und begründen Sie Ihre Ideen.

- Was bedeutet „auf Hochtouren laufen/arbeiten“? Wird diese Redewendung positiv oder negativ betrachtet?
- Kennen Sie andere Redewendungen, die mit Arbeit zusammenhängen?

arbeiten wie ein _____

beide _____ *voll zu tun haben*

_____ *wie die Biene sein*

Aufg. 9: Ergänzen Sie anhand von der Merkmalskarte das entsprechende Verb.

Anfangen oder ***beginnen***?

Lösung:

Die Tagung _____ am Freitag um 14.30 Uhr.

beginnt

Keiner weiß mehr, wann es _____, wann es aufhört.

anfängt

Mit 14 Jahren habe ich _____ im Chor zu singen.

angefangen

Im Herbst dieses Jahres _____ ein neuer Ausbildungskurs.

beginnt

Marie sah aus, als würde sie gleich _____ zu heulen.

anfangen

Ich habe sogar _____ Deutsch zu lernen.

angefangen

Die Vorstellungen _____ um 20.15 Uhr.	beginnen
Wir müssen _____, darüber zu reden.	anfangen
Ich wusste gar nicht, wo ich _____ sollte.	anfangen
Der Bau eines neuen Hafens _____ am 1. Oktober.	beginnt
Am Wochenende _____ in den Stadtparks die Saison.	beginnt

Aufg. 10: Tragen Sie in die Tabelle ein, welche Kontextbereiche zu den Verben *anfangen* und *beginnen* gehören.

BEGINNEN	ANFANGEN
WETTBEWERB	GESELLSCHAFTLICHE EREIGNISSE

➤ In welchen Kontexten können beide Verben verwendet werden?

7. Fazit

Das Vorhaben meiner Arbeit war es, mithilfe der strukturierten lexikalischen Merkmalskarten die unterschiedlichen Verwendungsweisen zwischen nahen Synonymen festzustellen. Es wurden solche Synonyme untersucht, die sich in die tschechische Sprache lediglich mit einem Äquivalent übersetzen lassen. In der ersten Phase der Arbeit habe ich mich deshalb mit der Synonymie aus theoretischer Sicht beschäftigt. Es wurde darauf aufmerksam gemacht, dass die Bedeutung dieses Begriffs in der Lexikologie nicht problemlos empfunden wird. Linguisten fassen die Definition und Klassifikation der Synonymie auf verschiedene Weise auf. Besonders umstritten ist der Terminus *absolute Synonymie*, die es nach vielen Lexikologen in der Praxis nicht gibt, denn nach der Meinung vieler Linguisten existieren kaum zwei Lexeme mit einer ganz identischen Bedeutung bzw. Verwendung. Daraus folgt, dass auch die von mir ausgewählten Synonyme semantische, stilistische oder syntaktische Unterschiede aufweisen. Beim Blick auf die CNS-Merkmalskarten ist das tatsächlich zu beobachten.

Die Merkmalskarten werden durch den Algorithmus der Kookkurrenzdatenbank CCDB strukturiert und ausgearbeitet. Die CCDB ist ein Bestandteil des Deutschen Referenzkorpus, darum habe ich der Korpuslinguistik ein Kapitel gewidmet, wobei ich sie als eine strukturalistische Methodologie vorgestellt habe. Da die lexikalischen Merkmalskarten auf Grund der Kookkurrenzanalyse entstehen und organisiert werden, musste ich wichtige Begriffe erläutern, die mit dem lokalen Kontext (d. h. Kookkurrenz) zusammenhängen. An demonstrativen Beispielen habe ich erklärt, was unter den Begriffen „Kookkurrenzanalyse“ und „Kookkurrenzprofil“ zu verstehen ist. Dann wurden zwei Typen der Merkmalskarten präsentiert – die sog. SOM- und CNS-Merkmalskarten, mit deren Analyse ich die Untersuchung der nahen Synonyme im empirischen Teil der Arbeit durchgeführt habe. Es wurde gezeigt, auf welchem Prinzip die Merkmalskarten aufgebaut werden und wie sie analysiert werden können.

Im Fokus der ersten Analyse waren Verben *anfangen*, *beginnen* und *starten*. Trotz der scheinbaren semantischen Gleichheit der Verben *anfangen* und *beginnen* wurde festgestellt, dass diese Verben in den meisten Kontexten unterschiedlich verwendet werden. Dasselbe gilt auch für das Lehnwort *starten*, das im Hinblick auf *anfangen* und *beginnen* eine sehr spezifische Bedeutung hat. Die nächste zu analysierende

Synonymreihe bilden die Verben *herstellen*, *erzeugen* und *produzieren*. Die Kookkurrenzprofile dieser drei Verben unterscheiden sich nicht so markant voneinander, wie das bei der ersten synonymischen Reihe der Fall war. Das heißt, diese drei Verben sind deutlich synonymisch und besitzen nur einige feine Spezifika.

Im letzten Teil der Arbeit habe ich mich der Anwendung von der CCDB und den Merkmalskarten in der Praxis gewidmet. Die CCDB ist vor allem für Lexikografen von wesentlichem Nutzen. Sie ist ein nützliches Hilfsmittel bei der Entstehung eines Wörterbuchs und hilft Unklarheiten der Bedeutungen und Verwendungsweisen der Lexeme aufzuklären. Eine andere Möglichkeit der praktischen Anwendung ist der Gebrauch im Sprachunterricht (Deutsch als Fremdsprache) für fortgeschrittene Lernende. Die Merkmalskarten können bei der Aneignung neuer Wörter, beim mündlichen oder schriftlichen Ausdruck oder auch im Unterricht der Landeskunde hilfreich sein. Als anschauliches Beispiel habe ich meinen eigenen Entwurf eines möglichen Arbeitsblattes präsentiert, der anhand der CNS-Merkmalskarte *anfangen* – *beginnen* ausgearbeitet wurde.

Zusammenfassend ist festzuhalten, dass diese Bachelorarbeit aufgeklärt hat, welche Aufgabe die Korpuslinguistik und vor allem die CCDB bei sprachwissenschaftlichen Untersuchungen erfüllen. Die CCDB stellt eine wichtige Alternative nicht nur für Linguisten und Lexikografen, sondern auch für Deutschlehrer dar.

8. Literaturverzeichnis

Wörterbücher:

Duden (2011): *Deutsche Universalwörterbuch* [CD-ROM]. 7., überarbeitete und erweiterte Auflage. Mannheim.

Fin Publishing (2005): *Německo-český a česko-německý studijní slovník*. 4. doplňné vyd. Olomouc.

Wissenschaftliche Studien:

Belica, Cyril (2007): *Kookkurrenzdatenbank CCDB - V3. Eine korpuslinguistische Denk- und Experimentierplattform für die Erforschung und theoretische Begründung von systemisch-strukturellen Eigenschaften von Kohäsionsrelationen zwischen den Konstituenten des Sprachgebrauchs*. <http://corpora.ids-mannheim.de/ccdb/>.

Bußmann, Hadumod (2002): *Lexikon der Sprachwissenschaft*. 3., aktual. und erweit. Aufl. Stuttgart.

Cruse, Alan (2011): *Meaning in language*. Oxford

Čermák, František: *Korpusová lingvistika Praha 2011*. Praha

Čermák, František / Filipec, Josef (1985): *Česká lexikologie*. 1. vyd. Praha.

Filipec, Josef (1961): *Česká synonyma z hlediska stylistiky a lexikologie*. Vyd. 1. Praha.

Heusinger, Siegfried (2004): *Die Lexik der deutschen Gegenwartssprache*. München.

Keibel, Holger / Belica, Cyril (2007): *CCDB: A Corpus-Linguistic Research and Development Workbench*. In: *Proceedings of Corpus Linguistics 2007*, Birmingham. http://corpus.bham.ac.uk/corplingproceedings07/paper/134_Paper.pdf

Kohonen, Teuvo (1995): *Self-Organizing Maps*. Berlin.

Linke, Angelika / Nussbaumer, Markus / R. Portmann, Paul (2004): *Studienbuch Linguistik*. 5., erweiterte Aufl. Tübingen.

Marková, Věra (2012a): *Synonyme unter dem Mikroskop*. Eine korpuslinguistische Studie. Tübingen.

Marková, Věra (2012b): *Zur neuen Bestimmung von synonymischen Reihen mit Hilfe der korpuslinguistischen Tools (self-organizing maps, contrasting near-synonyms)*. In: *Acta Universitatis Carolinae Philologica 2 / Germanistica Pragensia XXII*. Praha, 59-66.

McEnery, Tony / Hardie, Andrew (2012): *Corpus linguistics: method, theory and practice*. Cambridge.

Perkuhn, Rainer / Belica, Cyril (2006): *Korpuslinguistik - das unbekannte Wesen oder Mythen über Korpora und Korpuslinguistik*. In: Sprachreport 1; 2-6.

Perkuhn, Rainer / Keibel, Horger / Kupietz, Marc (2012): *Korpuslinguistik*. Paderborn.

Schippan, Thea (2002): *Lexikologie der deutschen Gegenwartssprache*. 2., unverändr. Aufl. Tübingen.

Schmidt, Marek (2012): *Suche nach neuen Wegen der Bedeutungsvermittlung von nahen Synonymen im fortgeschrittenen Fremdsprachenunterricht – eine unendliche oder endlose Geschichte?*. In: Acta Universitatis Carolinae Philologica 2 / Germanistica XXII. Praha, 123-131.

Storch, Günther (1999): *Deutsch als Fremdsprache: theoretische Grundlagen und praktische Unterrichtsgestaltung*. München.

Storjohan, Petra (2006): *Kontextuelle Variabilität synonymer Relationen*. In: OPAL - Online publizierte Arbeiten zur Linguistik 1/2006. Mannheim, 1-24.

Tognini-Bonelli, Elena (2001): *Corpus Linguistic at Work*. Amsterdam.

Ullmann, Stephen (1967): *Grundzüge der Semantik: die Bedeutung in Sprachwissenschaftlicher Sicht*. Berlin.

Vachková, Marie (2007a): „...und jedes Wort hat fließende Grenzen...“ – Gedenkschrift für Prof. PhDr. Alena Šimečková, CSc. Praha, 41-63.

Vachková, Marie (2007b): *Kapitoly k německo-české metalexikografii I*. Praha.

Vachková, Marie / Schmidt, Marek / Belica, Cyril (2007): *Prager Wanderungen durch die Mannheimer Quadrate*. In: Sprachreport, Sonderheft März 2007. Mannheim, 16-21.

Vachková, Marie (2008): *Zur Erforschung und Erfassung der diskursgebundenen semantischen Kontraste auf der Grundlage des SOM-Modells*. In: Acta Universitatis Carolinae Philologica 2 / Germanistica XX. Praha, 193-207.

Vachková, Marie / Marková, Věra / Belica, Cyril (2008): *Korpusbasierte Wortschatzarbeit im Rahmen des fortgeschrittenen Germanistikunterrichts*. In: Zielsprache Deutsch, 20-35.

Vachková, Marie / Belica, Cyril (2009): *Self-Organizing Lexical Feature Maps. Semiotic Interpretation and Possible Application in Lexicography*. In: IJGLSA 13, 2. Berkeley, 223-260. <http://corpora.ids-mannheim.de/IJGLSA.pdf>

Internetquellen:

<http://corpora.ids-mannheim.de/ccdb/>, Stand 1/2014

<http://www1.ids-mannheim.de/kl/projekte/korpora/>, Stand: 1/2014

9. Im Text verwendete Abkürzungen

CCDB	Co-occurenc Database
CNS	Contrasting Near-Synonyms
DeReKo	Das Deutsche Referenzkorpus
DUW	Duden Universalwörterbuch
KP(e)	Kookkurrenzprofil(e)
MK(en)	Merkmalskarte(n)
SOM	Self-Organizing Map